DATA-DRIVEN OPTIMIZATION IN POWER SYSTEMS OPERATIONS

By

ALI BAGHERI

Bachelor of Science in Applied Mathematics
Yazd University
Yazd, Iran
2005

Master of Science in Industrial Engineering
Mazandaran University of Science and Technology
Babol, Iran
2008

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
June, 2018

DATA-DRIVEN OPTIMIZATION IN POWER SYSTEMS OPERATIONS

Dissertation Approved:

Dr. Chaoyue Zhao

Dissertation Advisor

Dr. Tieming Liu

Dr. Farzad Yousefian

Dr. Yuanxiong Guo

Dedicated to my

beloved parents, Azam and Hossein,

beloved wife, Farideh,

and beloved daughter, Diana

The dedication reflects the views of the author and are not endorsed by committee members or Oklahoma State University.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest appreciation to my adviser, Dr. Chaoyue Zhao, who helped me a lot during my Ph.D. Study. I will be forever thankful for her mentorship, support and encouragement, and for everything she has taught me without which I would not have reached this point. She has not only been a great adviser, but also a supportive and caring friend. I am deeply thankful for her patience and kindness towards me and my family during difficulties in life.

I am also very grateful to my committee members, Dr. Tieming Liu, Dr. Farzad Yousefian, and Dr. Yuanxiong Guo for their valuable guidance and supports during my Ph.D. process. I also would like to really thank Dr. Jianhui Wang for his constructive guidance and supports in my research works.

I would like to sincerely thank the department of Industrial Engineering and Management (IEM) at Oklahoma State University (OSU). I am especially thankful to Prof. Sunderesh Heragu, head of the department, for all his supports during my Ph.D. and job search. I would also like to thank Prof. Manjunath Kamath, the former graduate coordinator, Dr. Balabhaskar Balasundaram, graduate coordinator, for their helpful advices and suggestions. My warm thanks go to all administrative staff at IEM, especially Laura Brown, for all worthwhile and sincere efforts.

I sincerely appreciate the financial support by grants from the National Science Foundation (ECCS-1610935 and CMMI-1662589).

My heartfelt thanks to my mother and father for their boundless love and all

Acknowledgments reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

sacrifices they made to get me where I am now. I would also like to thank my brother and sister for all their support and encouragement.

Last but not least, I wish to heartily thank my lovely wife, Farideh. Without her unconditional support, patience and encouragement, I would not have been able to finish this journey successfully. I also owe a heartfelt thank to my daughter, Diana, for her presence in my life, who encourages me to work harder.

Acknowledgments reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: Ali Bagheri

Date of Degree: June, 2018

Title of Study: DATA-DRIVEN OPTIMIZATION IN POWER SYSTEMS OPERATIONS

Major Field: Industrial Engineering and Management

Today's power systems are large scale systems consisting of multiple generating stations, load zones (distributors or utilities) and very complex interconnected power transmission networks. One of the major issues for power systems operators is that they face with several sources of uncertainty such as equipment failure uncertainty and demand uncertainty in the power systems operations. Moreover, the growing trend in renewable energy capacity installments has added a higher level of uncertainty to power systems operations. Therefore, uncertainty management has recently become one of the most challenging issues in power system operations and control.

However, in many cases, partial information of the uncertain parameters are available. Distributionally robust optimization is a newly emerged optimization approach to address optimization problems under uncertainty with partial information. In this study, we develop efficient distributionally robust optimization models to address several challenging problems arising in power systems operations.

First, we propose a data-driven approach to solve the stochastic transmission expansion planning problem under demand uncertainty. Then, we develop a data-driven approach to deal with the stochastic transmission system hardening planning problem in the presence of wind generation uncertainty and multiple simultaneous disruptive events. Afterward, we propose two reliability analysis schemes for the power transmission system hardening under distributional uncertainty of random contingencies. Finally, we present a data-driven chance-constrained stochastic unit commitment (power generation scheduling) under wind power uncertainty, in which the chance constraint controls and limits the level of energy imbalance. In all cases, we reformulate the original problems to two-stage stochastic mixed integer programs. Then, we deploy decomposition approaches to solve the developed models.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

An electric power system is a network consisting of three different systems as its main components: (1) generation system, (2) transmission system and (3) distribution system. Also, an electricity market is a system, in which market participants trade energy through bidding purchase prices by buyers and offering electricity prices by sellers. In the United States, from the past two decades, the structure of electricity markets has been changing from the traditional regulated model of vertically integrated electric utilities to a new structure i.e., deregulated electricity market [21]. In the former, each utility has its own generation units (suppliers) and is responsible for serving its own customers, while in the latter, the electric industry is decomposed into the three above-mentioned systems (generation, transmission and distribution) which are operated independently in a competitive market. The most likely expected advantage of deregulated electricity market is the electricity price reduction due to the fact that energy price is no longer regulated and customers are not obliged to buy electricity from a specific supplier in the same region. They are able to buy energy from suppliers in other regions with competitive prices [96].

Changing the energy market structure has required an independent operational control of the power grid. In this regard, Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs), as non-profit independent organizations, have been established to coordinate, manage and control the operations of market participants through making necessary standards and rules. For brevity, hereinafter we use "ISOs" to refer both ISOs and RTOs. In most cases, ISOs take control

of power systems established across multiple US states. ISOs must be independent of every market participants such as generation units, transmission companies, electric utilities, end-use customers and etc. They also should be open access and act on a non-discriminatory basis to all market participants using well expanded transmission systems [96]. Currently, there are nine ISOs in the North America, of which seven are in the US. The first ISO, California ISO (CAISO), currently serving the most of California and a portion of Nevada, established in 1996. Also, Midcontinent ISO (MISO) is the widest ISO coordinating all or portions of 15 states of the US (available at https://www.ferc.gov/industries/electric/indus-act/rto.asp).

Restructuring the electricity market has resulted in a creation of large scale power systems consisting of multiple generating stations, multiple load zones (distributors or utilities) and very complex interconnected power transmission networks. ISOs are responsible for maintaining the reliability of such complex power systems. According to North American Electric Reliability Council (NERC), reliability is the degree of power system performance under which customer's electricity demand is supplied and is delivered under the accepted standards [95]. This definition of reliability contains two concepts: (1) Adequacy: the ability of a power system to supply the customers electricity demand via available generation units and transmission systems and reserves. (2) Security: the ability of a power system to keep working after some contingencies such as transmission line outages or equipment failures. In reality, each ISO communicates with and coordinates several balancing authorities in its covered region. Balancing authorities are responsible for minutely maintaining the system reliability through keeping the supply and demand balance within their borders. There may be none, one or multiple generators under the control of each balancing authority and also an aggregated electricity demand assigned to it, i.e. load. In the literature of power systems research, balancing authorities are referred to as buses.

In order to maintain the reliability of such complex and uncertainty-integrated

power systems, ISOs run a day-ahead Unit Commitment (UC) with objectives such as minimizing the overall operational cost or maximizing the social welfare while satisfying all system constraints across the covered region [128, 122]. Also, ISOs conduct Generation and Transmission Expansion Planning (GTEP) to satisfy electricity demand and to maximize electricity trade opportunities in their markets [26]. In addition, to improve power system resilience against possible disruptions caused by catastrophic events, ISOs conduct Transmission Hardening Planning (THP) [74, 109].

In addition, one of the major issues in power system operations is that they face with two types of uncertainties: (1) equipment failure uncertainty and (2) forecasting uncertainty [91]. The former exists because of either unreliable nature of equipments or disruptive events such as natural disasters, terrorist attacks and etc, while the latter comes from the errors in forecasting practices such as electricity demand forecasting, renewable energy generation forecasting and etc. Moreover, the growing trend in renewable energy capacity installments has added higher levels of uncertainty to power system operations. Therefore, uncertainty management has recently become one of the most challenging issues in power system operations and control [104]. To cope with uncertainties, there are two remedies proposed in the power systems literature. One is considering reserve requirement and the other is developing and implementing optimization-under-uncertainty models [91]. However, the more wind capacity is integrated into power systems, the more uncertainty they would be confronted with, so the more reserve capacity would be required to cope with forecasting errors [38]. For example, in Texas in February 2008, a wind generation ramp down of 1700 MW occurred within three and a half hours and caused a significant curtailment of industrial demand [83]. Therefore, to maintain power grid reliability, the application of optimization-under-uncertainty modeling techniques has received more attention.

Recently, stochastic programming (SP) [97, 36, 107] and robust optimization (RO) [13, 128, 52] approaches have been successfully employed to cope with power systems

operations under uncertainty. However, there are some disadvantages with both SP and RO. In SP, it is assumed that the random parameter follows a certain probability distribution. However, the accurate probability distribution is very hard to obtain and inaccurate distribution assumptions may lead to unreliable decisions. Also, robust optimization approach, which is based on the worst-case scenario of the random parameter, is too conservative and costly inefficient, because the worst-case scenario occurs rarely. To address the drawbacks of stochastic programming and robust optimization approaches, distributionally robust / data-driven optimization (DRO) approaches (see, e.g., [37, 46, 55] among others) have been recently developed.

In DRO approaches, instead of considering a particular probability distribution for the uncertain parameter, as SP does, an unknown and ambiguous probability distribution is considered. Also, instead of considering the worst-case scenario of the uncertain parameter, as RO does, the worst-case probability distribution is considered. To this end, by learning from a set of historical data or moment information of the random parameter, an ambiguity set for the unknown probability distribution of the random parameter is constructed. Then, the objective is to minimize the total cost or to maximize the total profit under the worst-case distribution scenario within the constructed ambiguity set. DRO is a risk-averse approach and is still conservative because it considers the worst-case distribution. However, in general DRO is less conservative than RO [46] .

Accordingly, in this research, due to the availability of considerable amount of historical data (such as demand data, renewable energy generation data, etc) to ISOs, we develop DRO models to address above-mentioned problems in power system operations under uncertainty.

In chapter 2, we propose a data-driven approach to solve the stochastic transmission expansion planning problem under demand uncertainty. In chapter 3, we develop a data-driven approach to deal with the stochastic transmission system hardening

planning problem in the presence of wind generation uncertainty and multiple simultaneous disruptive events. In chapter 4, we propose two reliability analysis schemes for the power transmission system hardening under distributional uncertainty of random contingencies. In chapter 5, we present a data-driven chance-constrained stochastic unit commitment, in which the chance constraint controls the level of energy imbalance. In chapter 6, we conclude this research.

CHAPTER 2

TRANSMISSION EXPANSION PLANNING VIA DATA-DRIVEN OPTIMIZATION

Due to the significant improvements of power generation technologies and the trend of replacing traditional power plants with renewable generation resources, the generation portfolio will experience dramatic changes in the near future. The uncertainty and variability of renewable energy and their sitting call for strategic and economic plans for expanding the transmission capacities. In this study, we develop a data-driven two-stage stochastic transmission expansion planning with uncertainties. In the proposed approach, purely by learning from the historical data, we first construct a confidence set for the unknown distribution of the uncertain parameters. Then, we develop a two-stage data-driven transmission expansion framework, by considering the worst-case distribution within the constructed confidence set, so as to provide a reliable while economic transmission planning decision. Furthermore, to tackle the model complexity, we propose a decomposition framework embedded with Benders' and Column-and-Constraint generation methods.

2.1 Problem Description and Literature Review

Transmission expansion planning (TEP) is the problem of how to expand the transmission network while satisfying the forecasted demand at minimum expansion and operational costs over a given planning horizon. Through conducting transmission expansion planning, ISOs can maximize electricity trade opportunities in their markets [26, 3, 43]. However, uncertainties always inherently exist in the transmission

expansion planning, such as regular uncertainties from demand and renewable energy, and occasional changes about new policies, generator retirement, etc. [92]. To address the uncertainties in TEP, stochastic programming (SP) and robust optimization (RO) approaches have been extensively employed, among others such as fuzzy models [25, 32, 27].

For stochastic programming, the uncertain parameters are either characterized by a limited number of scenarios as their possible realizations or assumed to follow a particular distribution. Then, the objective is to minimize the total expected cost or to maximize the expected profit or social welfare corresponding to the generated scenarios or assumed distribution. For instance, in [36], the uncertain demand is interpreted as a number of scenarios with the objective of maximizing the aggregate social welfare, while in [89], both transmission and generation expansion are considered under random outages of generation units and transmission lines as well as demand uncertainty. Recently, two-stage stochastic programs have been successfully used for transmission expansion problems. For example, in [58], a two-stage stochastic programming framework is proposed to address the generation and transmission expansion planning problem under equipment failure and load uncertainties, in which the first stage considers the expansion decisions and the second stage considers the operational costs and system reliability. Also, chance-constrained stochastic optimization approaches have been recently utilized in solving the TEP problems, in which chance constraints are utilized to enforce the probability of generation and transmission line capacity violations to be no more than a predefined level [116], [68].

Moreover, robust optimization approaches have been applied to many operational and planning problems arising in power systems, for instance, unit commitment [128, 56, 122, 14] and optimal bidding strategy [11, 41]. As for the TEP, robust optimization has also been used by taking different sources of uncertainty and different objective functions into account. For example, in [78], random contingencies

are considered as the source of uncertainty in a multi-stage framework, while in [52], intermittent renewable energy as well as the demand uncertainty are considered in a two-stage model. In addition, in [4], the investment cost of new transmission lines and demand information are considered as unknown parameters. In [29], with the objectives of minimizing the worst-case cost and the worst-case regret, a trilevel optimization model is proposed under the unknown generation expansion behavior of the electricity suppliers and the demand uncertainties. In general, for robust optimization approaches, to characterize possible realizations of the uncertain parameter, a deterministic uncertainty set, such as an interval, a cardinality uncertainty set, or a polyhedral uncertainty set, is constructed. The objective of the robust optimization based TEP framework is to minimize expansion and operational costs or regret by considering the worst-case scenario of random parameters in the predefined uncertainty set. Therefore, the robust optimization based approaches can provide reliable and conservative decisions.

However, both stochastic and robust optimization approaches face challenges in practice. For the stochastic optimization approach, the distributions of the uncertain parameters, such as renewable generation output, are often assumed known. However, the obtained scheduling can be sensitive to the distributions and thus can be biased in practice with an inaccurate distribution assumption. In this case, both cost effectiveness and system reliability can be compromised. On the other hand, for the robust optimization approach, the system performance is determined by considering the worst-case scenario of the random parameter, therefore, the solutions are often too conservative and pessimistic, which may lead to large system costs. Also, this approach requires limited information to construct the uncertainty set, which does not utilize the available historical data to a largest extent.

Many works have been done recently to address the above-mentioned challenges arising in stochastic programming and robust optimization approaches. For instance,

recent application of robust optimization approach uses a parameter called "budget of uncertainty" to control the level of conservativeness (see, e.g., [128, 56, 14, 11, 41, 52], among others). In addition, hybrid stochastic programming and robust optimization approaches have been recently proposed against over-conservativeness and heavy computational burden (see, e.g., [122, 39, 82, 44], among others). However, these approaches still consider the worst-case scenario of the random parameter in the corresponding uncertainty set in their objective functions, and they are not able to utilize historical data to a large extent to cope with uncertainty. To address these challenges, accordingly, distributionally-robust and data-driven optimization concepts have also been proposed.

For the distributionally-robust and data-driven optimization approaches, instead of considering the randomness of uncertain parameters, an unknown probability distribution is considered and characterized by learning from the available historical data [37, 46, 55, 125, 124]. Other than assuming any particular probability distribution of the random parameters as the traditional stochastic optimization approach does, a confidence set is constructed, with a certain confidence level (for example 95%) to cover the unknown probability distribution, and the probability distribution can run adversely within the confidence set. The objective of the model is to minimize the worst-case cost associated with the worst-case distribution in the confidence set. Though this approach is still a risk-averse approach, the conservativeness is in general less than that of the traditional robust optimization approach [46]. It is also shown that the level of conservativeness decreases as the size of historical data increases. The theoretical frameworks of the distributionally-robust and data-driven optimization approaches have been recently applied to power system optimization problems. For example, in [126], a data-driven two-stage stochastic program is utilized to address unit commitment under wind power uncertainty. In [15], a distributionally-robust optimization framework is presented to optimize reserve scheduling with partial infor-

mation of renewable energy. In [111], a two-stage distributionally-robust optimization model is proposed for jointly optimizing energy and reserve under wind uncertainty. In [**?**], a distributionally-robust chance-constrained model is formulated to address optimal power flow under wind, load and reserve capacity uncertainties. However, to the best of our knowledge, this is the first work to apply distributionally-robust / data-driven optimization concept to solve transmission expansion problems.

In this study, we propose a data-driven risk-averse two-stage stochastic transmission expansion planning framework, in which the first stage deals with transmission line expansion decisions, while the second stage considers the worst-case operational cost (production cost and load shedding penalty cost) associated with the worst-case distribution of the random demand. Note here that our proposed model can be equivalently applied with renewable energy (e.g. wind power) uncertainty, which will be discussed in more detail in section 2.5. Moreover, we develop a novel decomposition approach that combines both Benders' and Column-and-Constraint generation methods [120] to address the problem. The contributions of this research can be listed as follows:

1. A data-driven risk-averse stochastic optimization framework is utilized to address the transmission expansion planning problem under uncertainties, which can provide a more robust transmission expansion decision than the traditional stochastic optimization approach, while a more economic expansion decision than the traditional robust optimization approach.

2. The proposed approach utilizes historical data information to a larger extent, as compared with robust optimization approaches. In addition, a reformulation of the proposed framework can be obtained, which does not lead to a heavier computational burden as more data is received. However, when the size of historical data becomes larger, and the conservativeness of the proposed approach

decreases and eventually vanishes as the size of data goes to infinity.

2. A new decomposition framework is proposed to solve the problem. The computational results show that the proposed decomposition method leads less computational time as compared with Benders' decomposition and Column-and-Constraint generation.

The remaining parts of this chapter are organized as follows: In Section 2.2, we define sets, parameters and variables. In Section 2.3, we construct the confidence set for the unknown probability distribution of the uncertain load. Then, the data-driven two-stage stochastic formulation is presented. In section 2.4, the proposed decomposition technique along with the solution algorithm are described. In section 2.5, the extensions are provided. In section 2.6, numerical results from the case studies are discussed. Finally, the research is concluded in section 2.7.

2.2 Nomenclature

A. Indices

$\mathcal{H}$	Index set of time periods in the planning horizon (e.g., 20 years).
$\mathcal{T}$	Index set of load blocks.
$\mathcal{B}$	Index set of all buses.
$\mathcal{E}_e$	Index set of existing transmission lines.
$\mathcal{E}_c$	Index set of candidate transmission lines.
$\mathcal{B}_i$	Index set of all buses directly connected to bus i.
$\mathcal{K}$	Index set of technology types.

B. Parameters

$V_{ij,t}^{T}$ Construction cost for new transmission line (i, j) at year t.

$V_{i,k,t}^{G}$ Investment cost for new generation unit of technology k at bus i at year t.

$W_{i,k,t}$ Power production cost of technology k at bus i at year t.

$L_{i,t}$ Load shedding cost at bus i at year t.

F_{ij} Flow capacity of transmission line (i, j).

$C_{i,k,t}^{g}$ Existing generation unit capacity of technology k at bus i at year t.

$C_{i,k,t}^{e}$ Candidate generation unit capacity of technology k at bus i at year t.

ν Market interest rate.

X_{ij} Reactance of transmission line (i, j).

θ_{i}^{min} Phase angle lower limit at bus i.

θ_{i}^{max} Phase angle upper limit at bus i.

$d_{i,t,b}$ Demand at bus i at year t load block b.

$w_{i,t,b}$ Renewable energy output at bus i at year t load block b.

C. Decision Variables

$z_{ij,t}$ Binary decision variable to indicate whether transmission line (i, j) is constructed at year t.

$y_{ij,t}$ Binary decision variable to indicate whether transmission line (i, j) exists at year t.

$u_{i,k,t}$ Binary decision variable to indicate whether new generation unit of technology k is constructed at bus i at year t.

$r_{i,k,t}$ Binary decision variable to indicate whether generation unit of technology k at bus i exists at year t.

$x_{i,k,t,b}$ Power generation of technology k at bus i at year t load block b.

$f_{ij,t,b}$ Power flow from bus i to bus j on transmission line (i,j) at year t load block b.

$\theta_{i,t,b}$ Phase angle at bus i at year t load block b.

$s_{i,t,b}$ Load shedding at bus i at year t load block b.

$w^c_{i,t,b}$ Renewable energy curtailment at bus i at year t load block b.

2.3 Problem Formulation

In this section, we propose a data-driven two-stage stochastic TEP model. The objective of the proposed model considers the construction cost in the first stage and the expected operational and load shedding costs corresponding to the real-time demand realizations in the second stage. In the model, the first-stage decisions include the construction decisions for all the candidate lines, before observing the true demand scenario. The second-stage decisions include the power generation level for each bus, power flow and phase angle for each transmission line, after knowing the actual demand. However, as we described in Section 2.1, it is biased to assume any particular distribution of the demand, which is usually unknown in practice. Instead, a set of historical data is available. In this study, we allow the probability distribution of the demand to be ambiguous and run adversely in a confidence set, and we learn from the historical data to construct the confidence set by introducing several distance measures of probability distributions. Note here that the historical data may not be

able to fully characterize the demand behavior in the future. Many other factors, such that smartgrids, electric vehicles, and demand response, may affect the future demand as well. However, in this study, we will focus on the utilization of historical data without considering the impact of these new technologies. Based on this, we develop a data-driven stochastic optimization approach to provide a risk-averse TEP decision with the demand uncertainty.

2.3.1 Confidence Set Construction

One approach to construct the confidence set is based on the moment information of the random parameter [37, 131]. By estimating the the mean values (denoted by $\hat{\mu}$) and the covariance matrices (denoted by Σ_0) of the uncertain parameters through learning from the historical data, the moment-based confidence set $\mathbb{D}$ can be constructed as:

$$\mathbb{D} = \big\{ P \in \mathcal{M}_+ : (E_P[\xi] - \hat{\mu})^T \Sigma_0^{-1} (E_P[\xi] - \hat{\mu}) \leq \epsilon_1, \\ E_P[(\xi - \hat{\mu})(\xi - \hat{\mu})^T] \preceq \epsilon_2 \Sigma_0 \big\}, \tag{2.1}$$

where $\mathcal{M}_+$ denotes the set of all distributions. Here, ϵ_1 and ϵ_2 can depend on the size of historical data. In this case, a larger size of historical data leads to lower values of ϵ_1 and ϵ_2; and accurate values of the first and second moments would be attainable as the size of data goes to infinity. However, even with fixed first and second moments, there are still an infinite number of probability distributions in $\mathbb{D}$.

In this research, we adopt a distribution-based approach to construct the confidence set $\mathbb{D}$. That is, instead of considering moment information, we construct the confidence set based on the distribution information. The distribution-based confidence set is shown as follows:

$$\mathbb{D} = \{ P \in \mathcal{M}_+ : d(P, \hat{P}) \leq \varphi \}, \tag{2.2}$$

where P denotes the true distribution, $\hat{P}$ represents the reference distribution derived from the historical data, $d(P, \hat{P})$ is the predefined probability distance between P and $\hat{P}$; and φ represents a tolerance level of the distance. To obtain the reference distribution, both parametric and nonparametric methods have already been utilized. Parametric estimation methods assume the random parameter follows a predefined distribution (e.g., normal distribution) and the parameters (e.g., the mean and variance) of the assumed distribution can be estimated through learning from the historical data [101]. Here, we adopt nonparametric estimation methods that can get rid of the distribution assumption.

2.3.2 Reference Distribution

We use the histogram as our reference distribution. That is, we partition the sample space Ω into N bins, i.e., $\Omega = \bigcup_{n=1}^{N} B_n$. Then, by counting the frequency of data samples falling into each bin, S_n, we can determine the reference distribution $\hat{P} = (\hat{p}_1, \hat{p}_2, ..., \hat{p}_N)$, in which $\hat{p}_n = S_n/S$, $\forall n = 1, 2, \cdots, N$. Specially, for the case that the true distribution is discrete, to get the reference distribution, for each scenario n, we count the number of historical data samples matching scenario n, and then divided by S to calculate the corresponding probability. Note here that the computational complexity of our data-driven approach only depends on the number of bins in the histogram. Therefore, the computational complexity of our proposed approach does not change as long as the number of bins is unchanged.

2.3.3 Probability Metrics and Value of φ

Intuitively, as we observe more information of the true distribution, i.e., more historical data samples, we can get a better estimation of the reference distribution, that is, the "distance" between the reference distribution and the true distribution becomes smaller. To measure the distance between two distributions, we can apply different

probability metrics. For instances, in [126], L_1 and L_∞ probability metrics have been applied to construct the confidence set:

$$d_1(P, \hat{P}) = \|P - \hat{P}\|_1 = \sum_{n=1}^{N} |p_n - \hat{p}_n|, \tag{2.3}$$

$$d_\infty(P, \hat{P}) = \|P - \hat{P}\|_\infty = \max_{1 \le n \le N} |p_n - \hat{p}_n|. \tag{2.4}$$

The Wasserstein metric has also been studied recently to construct the confidence set:

$$d_W(P, \hat{P}) = \inf_J \{E_J[d(X, Y)] | X \sim P,\ Y \sim \hat{P}\}, \tag{2.5}$$

where J denotes all joint distributions of the random variables X and Y with marginal distributions P and $\hat{P}$. Accordingly, different metrics lead to different convergence rates, i.e., the value of φ. For instance, for the above-introduced metrics, the values of φ are listed as follows, respectively:

$$\varphi_1 = (N/2S) \log(2N/1 - \beta), \tag{2.6}$$

$$\varphi_\infty = (1/2S) \log(2N/1 - \beta), \tag{2.7}$$

$$\varphi_W = D\sqrt{(2/S) \log(1/1 - \beta)}, \tag{2.8}$$

where D is the diameter of Ω. For more details, the readers can refer to [125] and [124]. In this study, we use L_1 norm to construct the confidence set as below:

$$\mathbb{D} = \{P \in \ \mathbb{R}_+^N | \ \ \sum_{n=1}^{N} |p_n - \hat{p}_n| \le \varphi_1\}. \tag{2.9}$$

Based on (2.6), the tolerance level φ_1 depends on the size of data samples and the confidence level β such that as the size of historical data S goes to infinity, the value of φ_1 goes to 0 and consequently, $\hat{P}$ converges to P.

2.3.4 Data-Driven TEP Framework

With the given confidence set of the ambiguous distribution of the uncertain demand, we formulate the data-driven stochastic TEP framework as follows:

$$\min \quad \sum_{(i,j)\in\mathcal{E}_c}\sum_{t\in\mathcal{H}} V_{ij,t}^T z_{ij,t} + \max_{P\in\mathbb{D}} E[Q(y,z,\xi)] \tag{2.10}$$

$$s.t. \quad -y_{ij,t-1} + y_{ij,t} - z_{ij,t} \le 0, \quad \forall t \in \mathcal{H}, \forall (i,j) \in \mathcal{E}_c, \tag{2.11}$$

$$y_{ij,t-1} - y_{ij,t} \le 0, \quad \forall t \in \mathcal{H}, \forall (i,j) \in \mathcal{E}_c, \tag{2.12}$$

$$y_{ij,t}, z_{ij,t} \in \{0,1\}, \tag{2.13}$$

where $Q(y,z,\xi)$ is equivalent to

$$\min \quad \sum_{i\in\mathcal{B}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} (1+\nu)^t (W_{i,k,t}x_{i,k,t,b}(\xi) + L_{i,t}s_{i,t,b}(\xi)) \tag{2.14}$$

$$s.t.$$

$$\sum_{k\in K} x_{i,k,t,b}(\xi) + \sum_{j\in\mathcal{B}_i(.,i)} f_{ji,t,b}(\xi) - \sum_{j\in\mathcal{B}_i(i,.)} f_{ij,t,b}(\xi) + s_{i,t,b}(\xi) = d_{i,t,b}(\xi), \quad \forall i \in \mathcal{B}, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.15}$$

$$x_{i,k,t,b}(\xi) \le C_{i,k,t}^g, \quad \forall i \in \mathcal{B}, \forall k \in \mathcal{K}, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.16}$$

$$-F_{ij} \le f_{ij,t,b}(\xi) \le F_{ij}, \quad \forall (i,j) \in \mathcal{E}_e, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.17}$$

$$-F_{ij}y_{ij,t} \le f_{ij,t,b}(\xi) \le F_{ij}y_{ij,t}, \quad \forall (i,j) \in \mathcal{E}_c, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.18}$$

$$f_{ij,t,b}(\xi) = (\theta_{i,t,b}(\xi) - \theta_{j,t,b}(\xi))/X_{ij}, \quad \forall (i,j) \in \mathcal{E}_e, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.19}$$

$$(\theta_{i,t,b}(\xi) - \theta_{j,t,b}(\xi))/X_{ij} - f_{ij,t,b}(\xi) + (1-y_{ij,t})M \ge 0, \quad \forall (i,j) \in \mathcal{E}_c, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.20}$$

$$(\theta_{i,t,b}(\xi) - \theta_{j,t,b}(\xi))/X_{ij} - f_{ij,t,b}(\xi) - (1-y_{ij,t})M \le 0, \quad \forall (i,j) \in \mathcal{E}_c, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.21}$$

$$\theta_i^{min} \le \theta_{i,t,b}(\xi) \le \theta_i^{max}, \forall i \in \mathcal{B}, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.22}$$

$$x_{i,k,t,b}(\xi), s_{i,t,b}(\xi) \ge 0, \tag{2.23}$$

where the objective function (3.10) is to minimize the total expansion and expected operational costs over a given planning horizon under the worst-case distribution in $\mathbb{D}$. Expansion cost is considered as total investments in transmission line constructions, while overall operational cost includes generation costs and load shedding costs. Constraints (2.11) and (2.12) indicate the transmission line construction status. Constraints (3.15) represent the power balance constraints. Constraints (3.16) enforce the generation capacity limit at each bus. Constraints (2.17) and (2.18) enforce the power flow limit for existing transmission lines and candidate transmission lines, respectively. Constraints (2.19) determine the power flow in terms of nodal phase angles of the existing transmission lines. Constraints (2.20) and (2.21) determine the power flow in terms of nodal phase angles of candidate lines, which are equivalent to $f_{ij,t,b}(\xi) = (\theta_{i,t,b}(\xi) - \theta_{j,t,b}(\xi))/X_{ij}y_{ij,t}$ by using the big-M method. Finally, constraints (3.17) enforce phase angle limits.

In the second-stage objective function, due to the independence of different scenarios ξ^n, we can interchange the second-stage minimization and summation (corresponding to the expectation term) operations. Hence, the objective function of the data-driven risk-averse two-stage stochastic model can be reformulated as follows:

$$\begin{aligned}\min & \sum_{(i,j)\in\mathcal{E}_c}\sum_{t\in\mathcal{H}} V_{ij,t}^T z_{ij,t} \\ & + \max_{P\in\mathbb{D}} \min \sum_{n=1}^{N}\sum_{i\in\mathcal{B}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} p_{i,t,b}^n (1+\nu)^t \big(W_{i,k,t}x_{i,k,t,b}(\xi^n) \\ & \qquad\qquad + L_{i,t}s_{i,t,b}(\xi^n)\big). \end{aligned} \tag{2.24}$$

2.4 Solution Methodology

In this section, we employ a decomposition algorithm framework, which utilizes both Benders' decomposition method and the Column-and-Constraint generation procedure [120], to solve the proposed data-driven model.

2.4.1 Second-Stage Reformulation

We dualize constraints (3.15) to (4.10) to obtain the following (SUB) problem:

$$\omega(y,z) =$$

$$\max \sum_{n=1}^{N} \Bigg(\sum_{i\in\mathcal{B}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} (d^n_{i,t,b}\gamma^n_{i,t,b} - \theta^{max}_i \bar{\rho}^n_{i,t,b} + \theta^{min}_i \tilde{\rho}^n_{i,t,b}) - \sum_{i\in\mathcal{B}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} C^g_{i,k,t}\lambda^n_{i,k,t,b} - \sum_{(i,j)\in\mathcal{E}_e}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} F_{ij}(\bar{\eta}^n_{ij,t,b} + \tilde{\eta}^n_{ij,t,b}) + \sum_{(i,j)\in\mathcal{E}_c}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} \big(- F_{ij}y_{ij,t}(\bar{\mu}^n_{ij,t,b} + \tilde{\mu}^n_{ij,t,b}) + X_{ij}(y_{ij,t}-1)M(\bar{\sigma}^n_{ij,t,b} + \tilde{\sigma}^n_{ij,t,b})\big) \Bigg) \tag{2.25}$$

s.t.

$$\gamma^n_{i,t,b} - \lambda^n_{i,k,t,b} - (1+\nu)^t W_{i,k,t} p^n_{i,t,b} \le 0, \quad \forall i, \forall k, \forall t, \forall b, \forall n \tag{2.26}$$

$$\gamma^n_{i,t,b} - \gamma^n_{j,t,b} - \bar{\eta}^n_{ij,t,b} + \tilde{\eta}^n_{ij,t,b} + X_{ij}\tau^n_{ij,t,b} = 0, \quad \forall (i,j) \in \mathcal{E}_e, \forall t, \forall b, \forall n \tag{2.27}$$

$$\gamma^n_{i,t,b} - \gamma^n_{j,t,b} - \bar{\mu}^n_{ij,t,b} + \tilde{\mu}^n_{ij,t,b} - X_{ij}\bar{\sigma}^n_{ij,t,b} + X_{ij}\tilde{\sigma}^n_{ij,t,b} = 0,$$
$$\forall (i,j) \in \mathcal{E}_c, \forall t, \forall b, \forall n \tag{2.28}$$

$$\sum_{(i,j)\in\mathcal{E}_e(.,i)} \tau^n_{ij,t,b} - \sum_{(i,j)\in\mathcal{E}_e(i,.)} \tau^n_{ij,t,b} + \sum_{(i,j)\in\mathcal{E}_c(i,.)} (\bar{\sigma}^n_{ij,t,b} - \tilde{\sigma}^n_{ij,t,b}) + \sum_{(i,j)\in\mathcal{E}_c(.,i)} (\tilde{\sigma}^n_{ij,t,b} - \bar{\sigma}^n_{ij,t,b}) - \bar{\rho}^n_{i,t,b} + \tilde{\rho}^n_{i,t,b} = 0, \quad \forall i, \forall t, \forall b, \forall n \tag{2.29}$$

$$\gamma^n_{i,t,b} - (1+\nu)^t L_{i,t} p^n_{i,t,b} \le 0, \quad \forall i, \forall t, \forall b, \forall n \tag{2.30}$$

$$p^n_{i,t,b} - k^n_{i,t,b} \le \hat{p}^n_{i,t,b}, \quad \forall i, \forall t, \forall b, \forall n \tag{2.31}$$

$$p^n_{i,t,b} + k^n_{i,t,b} \ge \hat{p}^n_{i,t,b}, \quad \forall i, \forall t, \forall b, \forall n \tag{2.32}$$

$$\sum_{n=1}^{N} k^n_{i,t,b} \le \varphi_1, \quad \forall i, \forall t, \forall b, \tag{2.33}$$

$$\sum_{n=1}^{N} p^n_{i,t,b} = 1, \quad \forall i, \forall t, \forall b \tag{2.34}$$

$$\bar{\eta}^n_{ij,t,b}, \tilde{\eta}^n_{ij,t,b}, \bar{\mu}^n_{ij,t,b}, \tilde{\mu}^n_{ij,t,b}, \bar{\sigma}^n_{ij,t,b}, \tilde{\sigma}^n_{ij,t,b}, \bar{\rho}^n_{i,t,b}, \tilde{\rho}^n_{i,t,b} \ge 0,$$

$$\lambda^n_{i,k,t,b}, p^n_{i,t,b}, k^n_{i,t,b} \ge 0, \gamma^n_{i,t,b}, \tau^n_{ij,t,b} \text{ are free,} \tag{2.35}$$

where $\gamma^n_{i,t,b}$ and $\lambda^n_{i,k,t,b}$ are dual variables of constraints (3.15) and (3.16), respectively, $\bar{\eta}^n_{ij,t,b}$ and $\tilde{\eta}^n_{ij,t,b}$ are dual variables of constraints (2.17), $\bar{\mu}^n_{ij,t,b}$ and $\tilde{\mu}^n_{ij,t,b}$ are dual variables of constraints (2.18), $\tau^n_{ij,t,b}$ is dual variable of constraint (2.19), $\bar{\sigma}^n_{ij,t,b}$ and $\tilde{\sigma}^n_{ij,t,b}$ are dual variables of constraints (2.20) and (2.21), respectively, $\bar{\rho}^n_{i,t,b}$ and $\tilde{\rho}^n_{i,t,b}$ are dual variables of constraints (3.17). In addition, constraints (2.31) - (2.33) represent the confidence set $\mathbb{D}$, where $p^n_{i,t,b}$ and $\hat{p}^n_{i,t,b}$ denote the true distribution and the reference distribution, respectively, with slack variable $k^n_{i,t,b}$ representing $|p^n_{i,t,b} - \hat{p}^n_{i,t,b}|$.

2.4.2 Cutting Planes

In the proposed decomposition framework, the following cuts are generated:

Benders' feasibility and optimality cuts

For the feasibility check, since we allow load shedding in the model, the second-stage constraints (3.15) to (4.10) are always feasible for any given first-stage variables y and z. Therefore, the first-stage feasibility is guaranteed and no feasibility check is needed.

As for optimality cuts, in each iteration, we obtain $\omega(y, z)$ for the given first-stage decisions y and z. Let ϑ represent the second-stage optimal objective value, we should have $\omega(y, z) \leq \vartheta$. Otherwise, if $\omega(y, z) > \vartheta$, the following optimality cut is added to the master problem:

$$\begin{aligned}
&\vartheta + \sum_{n=1}^{N} \Bigg(\sum_{(i,j)\in\mathcal{E}_c} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} F_{ij}(\bar{\mu}^n_{ij,t,b} + \tilde{\mu}^n_{ij,t,b}) y_{ij,t} - X_{ij} M(\bar{\sigma}^n_{ij,t,b} + \tilde{\sigma}^n_{ij,t,b}) y_{ij,t} \Bigg) \geq \\
&\sum_{n=1}^{N} \Bigg(\sum_{i\in\mathcal{B}} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} (d^n_{i,t,b}\gamma^n_{i,t,b} - \theta^{max}_i \bar{\rho}^n_{i,t,b} + \theta^{min}_i \tilde{\rho}^n_{i,t,b}) \\
&\quad - \sum_{i\in\mathcal{B}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} C^g_{i,k,t} \lambda^n_{i,k,t,b} - \sum_{(i,j)\in\mathcal{E}_e} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} F_{ij}(\bar{\eta}^n_{ij,t,b} + \tilde{\eta}^n_{ij,t,b}) \\
&\quad - \sum_{(i,j)\in\mathcal{E}_c} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} \big(X_{ij} M(\bar{\sigma}^n_{ij,t,b} + \tilde{\sigma}^n_{ij,t,b}) \big) \Bigg).
\end{aligned} \tag{2.36}$$

Column-and-Constraint Generation

According to the Column-and-Constraint generation method explained in [120], in iteration s, after we obtain the optimal distribution p^s from the (SUB) problem, if $\omega(y,z) > \vartheta$, then the following cut is added to the master problem:

$$\vartheta \geq \sum_{n=1}^{N} \sum_{i\in\mathcal{B}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{H}} \sum_{b\in\mathcal{T}} p_{i,t,b}^{ns}(1+\nu)^t \big(W_{i,k,t}x_{i,k,t,b}(\xi^n) + L_{i,t}s_{i,t,b}(\xi^n)\big). \quad (2.37)$$

2.4.3 Solution Algorithm

We have the following master problem (MP) for the proposed decomposition framework:

$$\begin{aligned} & \min \sum_{(i,j)\in\mathcal{E}_c} \sum_{t\in\mathcal{H}} V_{ij,t}^T z_{ij,t} + \vartheta \\ s.t.\quad & \text{Constraints } (2.11) - (3.11), \\ & \text{Constraints } (3.15) - (4.10) \text{ for each } \xi^n, \\ & \text{Constraints } (5.38) - (2.37), \end{aligned}$$

where cuts are added to the master problem in an iterative manner until the optimal solution or predefined optimality gap is achieved. The solution algorithm can be summarized in the following steps:

1. Initialization.
2. Solve MP and get the first-stage decision variables y, z and ϑ.
3. Fix y and z and solve the (SUB) problem and obtain $\omega(y,z)$.
4. Check the optimality condition. If $\omega(y,z) \leq \vartheta$, then stop and output the result. Otherwise, if $\omega(y,z) > \vartheta$, generate and add cuts (5.38) and (2.37) to MP and go to step 2.

2.5 Extensions and Future Work

This model can be easily extended to the case that renewable energy is another dimension of uncertainty in the power system. Let $w_{i,t,b}$ represent the renewable energy output and $w^c_{i,t,b}$ represent the curtailed amount of renewable energy, then we have:

$$\sum_{k\in K} x_{i,k,t,b} + \sum_{j\in\mathcal{B}_i} f_{ij,t,b}(\xi) + s_{i,t,b}(\xi) - w^c_{i,t,b}(\xi) = d_{i,t,b}(\xi) - w_{i,t,b}(\xi),$$
$$\forall i \in \mathcal{B}, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}, \tag{2.38}$$

$$w^c_{i,t,b}(\xi) \leq w_{i,t,b}(\xi), , \forall i \in \mathcal{B}, \forall t \in \mathcal{H}, \forall b \in \mathcal{T}. \tag{2.39}$$

In addition, the term $\sum_{i\in\mathcal{B}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} C_{i,t} w^c_{i,t,b}(\xi)$ is added to the second-stage objective function, where $C_{i,t}$ is the wind power curtailment penalty cost for bus i at time t. Note here that random parameter ξ follows a joint probability distribution of demand d and renewable energy output w.

In addition, the proposed data-driven framework can be extended to the joint transmission and generation expansion planning under demand uncertainty as follows:

$$\min \sum_{(i,j)\in\mathcal{E}_c}\sum_{t\in\mathcal{H}} V^T_{ij,t} z_{ij,t} + \sum_{i\in\mathcal{B}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{H}}\sum_{b\in\mathcal{T}} V^G_{ij,t} u_{i,k,t}$$
$$+(1+\lambda)^t (W_{i,k,t} x_{i,k,t,b} + L_{i,t} s_{i,t,b}) \tag{2.40}$$

$s.t.$

$$-r_{i,k,t-1} + r_{i,k,t} - u_{i,k,t} \leq 0, \quad \forall i, \forall t, \forall k, \tag{2.41}$$

$$r_{i,k,t-1} - r_{i,k,t} \leq 0, \forall i \in \mathcal{B}, \quad \forall i, \forall t, \forall k, \tag{2.42}$$

$$x_{i,k,t,b} \leq C^e_{i,k,t} r_{i,k,t} + C^g_{i,k,t}, \quad \forall i, \forall k, \forall t, \forall b, \tag{2.43}$$

Constraints (2.11) to (3.15), (2.17) to (4.10),

$$r_{i,k,t}, u_{i,k,t} \in \{0,1\}, \tag{2.44}$$

where constraints (2.41) and (2.42) indicate the new generation unit and existence, and constraint (2.43) indicates the generation capacity at each bus. Note here that

new generation installments can potentially affect the transmission expansion planning solutions. For example, if the generation expansion decisions are made in some buses, more generation capacities will be available in these buses to accommodate demand uncertainty, e.g., see constraint (2.43) as compared with constraint (3.16), which will potentially lead to less transmission expansions.

2.6 Case Study

In this section, to show the effectiveness of the proposed approach, we test a 6-bus system and an IEEE 118-bus system (available at http://motor.ece.iit.edu/data) to conduct numerical experiments. In this study, a planning horizon of 20 years is considered during which planning decisions are made yearly. In order to generate scenarios of the uncertain demand, we assume that demand follows a multivariate normal distribution and the mean values of demand increase by 1 percent yearly [75]. Also, we set the covariance to be 0.2 of its mean. Moreover, we consider three types of generation technologies (coal, natural gas and wind technologies) [29]. As it is mentioned in [75], it is predicted that incentive programs, (e.g. a tax credit program), which encourage investments in renewable and natural gas plants, will result in generation capacity increase by 1 percent annually in these types of plants. In addition, it is also predicted that the retirement plans for coal plants will lead to a yearly reduction by 1 percent in coal plant generation capacity. Hence, for numerical experiments, we increase the wind farm and natural gas plant capacities and decrease coal plant capacities by 1 percent annually. The market interest rate is assumed to be 0.1 yearly. To conduct the numerical experiments, we implement the proposed formulation and algorithm by C++ and CPLEX 12.6 and run it on a computer with Intel Xeon 3.2 GHz and 8 GB memory.

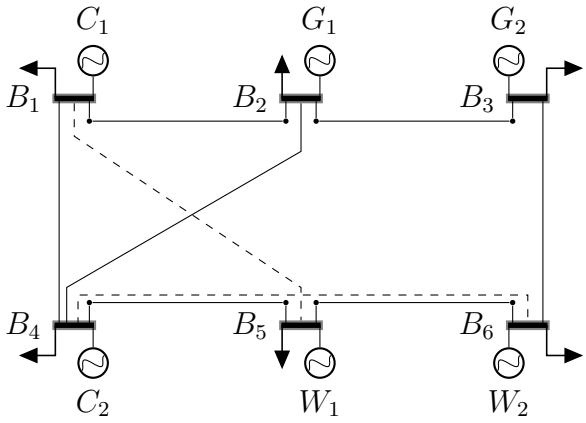

Figure 2.1: Modified 6-bus system

Table 2.1: 6-bus system modifications

Bus	Tech Type	Load (MW)	Generation Capacity (MW)	Generation Cost ($/MWh)	Load Shedding Cost (×1000$/MWh)
1	Coal	300	350	68	11.3
2	Gas	100	200	55	14.5
3	Gas	100	150	48	13.0
4	Coal	300	350	55	12.8
5	Wind	100	350	0	10.5
6	Wind	100	350	0	12.4

Table 2.2: 6-bus system candidate lines

Candidate Line	From	To	Reactance (Ω)	Capacity (MW)	Construction Cost($m)
1	1	5	0.016	500	224
2	4	6	0.037	500	110

2.6.1 6-Bus System

The 6-bus system, presented in Fig. 2.1, includes six buses and seven transmission lines. Specific settings for the 6-bus system are presented in Tables 2.1 and 2.2. We consider 2 coal plants, 2 natural gas plants and 2 wind farms, as well as 2 candidate transmission lines (denoted by dotted lines). In addition, the mean demands for each bus for the first year of the planning horizon are presented in the third column of Table 2.1.

For the 6-bus system, we first show the effects of the size of historical data and the confidence level on the conservativeness of the proposed approach. Then, we compare the performance of the proposed approach with the traditional two-stage stochastic optimization (e.g., [68]) and the traditional two-stage robust optimization (e.g., [52, 29]) approaches.

Table 2.3: Effects of the size of historical data on total cost

# of data		DDTEP		STEP		STEP_P	
	φ	Obj ($m)	CPU (s)	Obj ($m)	CPU (s)	Obj ($m)	CPU (s)
10	1.7269	698.8	2	595.9	2	623.2	2
50	0.3454	656.2	2	614.2	2	623.2	2
100	0.1727	642.1	2	613.2	2	623.2	2
500	0.0345	629.7	2	621.4	2	623.2	2
1000	0.0173	626.1	2	621.6	2	623.2	2
5000	0.0035	623.6	2	622.5	2	623.2	2
10000	0.0017	623.4	2	622.9	2	623.2	2

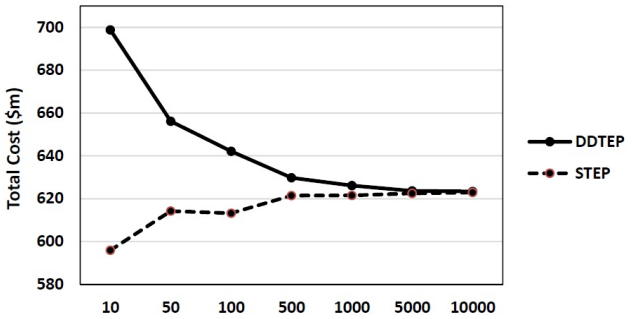

Figure 2.2: Effects of the size of historical data on total cost

Effects of the size of historical data

In this section, we investigate how the conservativeness level and total cost of the DDTEP model are affected by the size of historical data. We consider a range from 10 to 10000 for the number of historical data points and set the confidence level β to be 99%; then, we test the system performance of the data-driven stochastic transmission expansion planing model (denoted by DDTEP), the traditional two-stage stochastic transmission expansion planing with the estimated distribution learned from historical data (denoted by STEP) and with the perfect distribution information (denoted by STEP_P). For STEP_P, a set of 50000 historical data points is used to approximate the true probability distribution of the uncertain demand for each bus at each time period. It can be observed that as the size of historical data increases, the value of φ decreases. Accordingly, a smaller value of φ leads to a smaller confidence set which subsequently, leads DDTEP to be less conservative. Hence, as the size of historical data increases, the conservativeness of the model decreases and the objective value of DDTEP tends to decrease. It indicates that the more data system operators can utilize, the more money they can save. Table 2.3 and Fig. 5.1 show DDTEP converges to STEP_P as the number of historical data increases. That is, the conservativeness vanishes as we have a large number of data. In addition, with more historical data, STEP also converges to STEP_P, since more accurate estimation of the true distribution can be obtained with more data. We can also observe that DDTEP does not increase the

computational efforts as compared with STEP_P.

Table 2.4: Effects of the confidence level on total cost

β	DDTEP	
	φ	Obj($m)
0.5	0.0749	629.9
0.6	0.0805	630.8
0.7	0.0877	632.0
0.8	0.0978	633.5
0.9	0.1151	635.8
0.95	0.1325	638.0
0.99	0.1727	642.1

Effects of the confidence level

It can be observed from equality (2.6) that the construction of the confidence set depends on the value of confidence level β. Therefore, we conduct this experiment to illustrate the effect of confidence level β on the performance of DDTEP. We assume the number of historical data points to be 100 and the range of β to be from 0.5 to 0.99. As shown in Table 2.4, as the value of β increases, the value of φ_1 increases. A larger φ_1 can be inferred as having higher chance that the confidence set contains the unknown distribution. Therefore, with larger values of β, DDTEP becomes more conservative and its objective value increases. Indeed, this relationship provides an opportunity to system operators to adjust the conservativeness level of the system based on their preference of confidence level β.

Table 2.5: DDTEP versus STEP and RTEP

Approach	Plan	CC ($m)	WCD ($m)	SimD ($m)	CPU (s)
DDTEP	(2)(1)	110	696.5	675.3	2
RTEP	(1,2)(1,1)	334	701.8	684.5	2
STEP	-	0	699.5	676.8	2

Comparisons with stochastic and robust optimization approaches

In this section, to show the effectiveness of our proposed approach, we conduct numerical experiments on the 6-bus system to compare the performance of our DDTEP approach with the traditional two-stage stochastic transmission expansion planning (STEP) and the traditional two-stage robust transmission expansion planning (denoted by RTEP). We let the size of historical data to be 100 and the confidence level β to be 99%, and solve the transmission expansion planning problem under demand uncertainty using the DDTEP, STEP and RTEP, respectively. Then, we fix the first-stage expansion plans obtained from each approach and solve the second-stage problem for randomly simulated instances. We report the results in Table 2.5. The second and the third columns present the expansion plans and the associated construction costs, respectively. For example, the expansion plan obtained by DDTEP suggests candidate line 2 to be constructed in time period 1 and the expansion plan obtained by RTEP suggests both candidate lines 1 and 2 to be constructed in time period 1. However, STEP suggests no candidate line needs to be constructed, from which we can observe STEP is the least conservative model compared with DDTEP and RTEP. From Table 2.5 we can observe that, DDTEP performs better than both RTEP and STEP, since DDTEP results in lower total costs under both the worst-case distribution scenario (indicated by WCD) and the randomly simulated distribution scenario (indicated by SimD) in comparison with the other two. That is because on

one hand, as compared with DDTEP, STEP leads to no line constructions, so it is not able to accommodate large load fluctuations and therefore yields to more load shedding. On the other hand, as compared with DDTEP, RTEP is over conservative by sacrificing more average cost-effectiveness to consider the worst-case load realization, which happens rarely in practice.

2.6.2 118-Bus System

For the 118-bus system, we first compare DDTEP with STEP and RTEP. Then, we compare our proposed decomposition approach with Benders' decomposition and Column-and-Constraint generation approaches.

Comparisons with stochastic and robust optimization approaches

In order to demonstrate the effectiveness of the proposed approach, we also conduct numerical experiments on the modified IEEE 118-bus system. In this case, we allow the number of candidate lines to vary in the range of 9 to 13 and then we compare the proposed approach with the STEP and RTEP approaches. Also, we set the size of historical data to be 100 and the value of β to be 99%. We apply the same simulation procedure as the one previously and report the expansion plans in Table 2.6 for five different cases of candidate lines, obtained from the DDTEP, RETP and STEP.

In addition, the cost values corresponding to the expansion plans in Table 2.6 for both DDTEP and STEP are reported in Table 5.3. Although the expansion costs for DDTEP are higher than STEP, the proposed DDTEP results in lower total costs as compared with STEP for the whole planning horizon, under both WCD and SimD, for all five cases. That is because DDTEP leads to more new line constructions to hedge against the risk of electricity demand variability to avoid load shedding as compared with STEP, so it is a risk-averse approach. Therefore, in case of unexpected large load realization, DDTEP plan yields to less load shedding compared with STEP plan.

Table 2.6: Expansion plans

# of candidate lines	DDTEP	RTEP	STEP
9	(3,4,5,7,8) (1,1,1,5,1)	(1,3,5,8) (1,1,11,1)	(4,5,7) (1,1,1)
10	(3,4,5,7,8,10) (1,1,1,5,1,1)	(1,3,5,8,10) (1,1,11,1,1)	(4,5,7) (1,1,1)
11	(3,4,5,7,8,10,11) (1,1,1,5,1,1,1)	(1,3,5,8,10) (1,1,11,1,1)	(4,5,7,11) (1,1,1,1)
12	(3,4,5,7,8,10,11) (1,1,1,5,1,1,1)	(1,3,5,8,10) (1,1,11,1,1)	(4,5,7,11) (1,1,1,1)
13	(3,4,5,7,8,10,11) (1,1,1,5,1,1,1)	(3,5,8,10,13) (1,11,1,1,1)	(4,5,7,11) (1,1,1,1)

Table 2.7: DDTEP versus STEP

# of candidate lines	DDTEP			STEP		
	WCD ($m)	Sim_D ($m)	CPU (s)	WCD ($m)	Sim_D ($m)	CPU (s)
9	3597.9	2968.3	107	3611.5	2971.9	55
10	3598.9	2967.2	86	3611.3	2971.7	58
11	3596.4	2959.5	89	3607.9	2963.9	59
12	3596.4	2959.8	92	3608.3	2964.4	59
13	3597.4	2961.9	93	3609.6	2966.3	60

Consequently, DDTEP leads to less load shedding cost and therefore less total cost compared with STEP model. Furthermore, we compare the proposed DDTEP with

Table 2.8: DDTEP versus RTEP

# of candidate lines	DDTEP			RTEP		
	WCD ($m)	Sim_D ($m)	CPU (s)	WCD ($m)	Sim_D ($m)	CPU (s)
9	3597.9	2968.3	107	3600.6	3034.3	135
10	3598.9	2967.2	86	3604.3	3036.1	99
11	3596.4	2959.5	89	3601.2	3030.8	102
12	3596.4	2959.8	92	3601.1	3031.5	106
13	3597.4	2961.9	93	3602.2	2985.0	108

RTEP and report the computational results in Table 2.8. As compared with RTEP, DDTEP results in lower total costs under WCD and SimD. That is because RTEP is more conservative by considering the worst-case demand scenario and therefore leads to more system costs, as compared with DDTEP, in which the worst-case demand distribution is considered.

In general, by utilizing the historical data and allowing the ambiguous distribution, DDTEP leads to a more cost-efficient transmission expansion plan than the one of RTEP and a more reliable plan than the one of STEP. Indeed, considering the worst-case ambiguous distribution in the confidence set mitigates the effects of the blindly assumed distribution in STEP and lower utilization of historical information in RTEP. Consequently, the proposed data-driven approach leads to less load shedding for the real time, which results in higher reliability and cost efficiency.

Comparison with other separation algorithms

Here, we show the computational efficiency of our solution methodology (denoted as BCC) as compared with Benders' decomposition (BD) and Column-and-Constraint (CC) generation methods. We set the optimality gap to be 0.1%. In table 2.9, we

report the computational time for each approaches. It can be observed that CC works much better than BD, since in general CC leads to less iterations to achieve optima as compared with BD and therefore yields to less computational time. For more information of CC approach, interested readers can find in [120]. Moreover, Table 2.9 shows the proposed BCC outperforms both BD and CC methods in terms of computational time, since it generates both BD and CC cuts in the one iteration, which takes advantage of both BD and CC approaches.

Table 2.9: Comparing different separation methods

# of candidate lines	BCC (s)	CC (s)	BD (s)
9	107	125	>3600
10	86	126	>3600
11	89	141	>3600
12	92	142	>3600
13	93	142	>3600

2.7 Summary

In this chapter, we proposed a data-driven approach to solve the stochastic transmission expansion planning problem under demand uncertainty. By learning from the significant amount of historical data available for ISOs, our proposed approach considers the probability distribution of the random demand within a confidence set, instead of considering a particular probability distribution as traditional stochastic programming approaches do. The proposed approach is a risk-averse approach because it considers the worst-case probability distribution of the random parameter. However, as shown by numerical experiments, with the same level of confidence, our

approach can provide less conservative results with more historical data available; and theoretically, as the size of data goes to infinity the conservativeness ultimately disappears and our data-driven approach becomes risk-neutral. In addition, it is numerically shown that our proposed approach can achieve both the cost effectiveness and reliability, which bridges the gap between traditional stochastic and robust approaches. Also, the proposed decomposition method can decrease computational time comparing with Benders' decomposition and Column-and-Constraints generation, which therefore has potential to improve the computational efficiency for solving real-world large-scale problems. As a direction for future study, we will integrate recently emergent concepts such as smatgrids and electric vehicles into the data-driven transmission expansion model. Also, we will consider the real data set of demand and conduct experiments to see the performance of the proposed method under real data set.

CHAPTER 3

RESILIENT TRANSMISSION HARDENING PLANNING VIA DATA-DRIVEN OPTIMIZATION

Hardening components in transmission systems is a practice to improve system resilience against possible disturbances caused by natural disasters. In a power system with a very high penetration of renewable energy, the system hardening will be further complicated by the uncertainty and variability of renewable energy. In this work, we study the transmission line hardening planing problem in the context of probabilistic power flows injected by the high penetration of renewable energy. We assume that the probabilistic information of renewable energy is incomplete and ambiguous and propose a data-driven approach to approximate the renewable uncertainty sets. We then extend the $N-1$ security criteria to multiple simultaneous contingencies and seek for a hardening plan prepared for the worst-case scenarios. A two-stage data-driven stochastic model is formulated by considering the joint worst-case wind output distribution and transmission line contingencies. Then, we apply the Column-and-Constraints generation method to solve the proposed model. To test the effectiveness of the proposed approach, we conduct experiments on 24-bus and 118-bus test systems. We numerically show that the data-driven approach can effectively address the uncertainty ambiguity and the proposed approach can produce effective hardening plans that improve the system resilience.

3.1 Problem Description and Literature Review

Natural disasters, such as hurricane, wildfire, flooding, etc., become more frequent in the recent years [99]. Most often the natural disasters disrupt power system operations and cause service interruptions, ranging from short-term service losses to large-area extended outages. For example, according to a report by the Presidents Council of Economic Advisers and the U.S. Department of Energy [69], power outages caused by severe weather conditions constitute 58% of all outages and 87% of outages affecting 50000 or more electricity customers, during the period of 2003 to 2012. It is estimated that weather incurred outages have caused U.S. 60 billion USD annually [69]. These facts indicate the urgency of improving power system resilience against extreme weather events. To mitigate the natural disaster related risks and improve power grid resilience, many research works along with optimization tools focus on three main thrusts in the process of the power grid in reacting to the natural disasters: 1) pre-disaster system hardening and investment (e.g., [93, 90, 79, 7, 115, 23, 16, 118, 2, 117]), 2) emergency responses and corrective actions during or right after disasters (e.g., [105, 42, 100]), and 3) self-healing, rapid system restoration and damage assessment after disasters(e.g., [1, 67, 47, 119]). Hardening, as one of the most effective activities to increase the power system resilience, is defined as any physical change (such as undergrounding power lines, vegetation management, pole reinforcing, etc) to the power system infrastructure to make it less vulnerable to be damaged from severe weather conditions [74, 109].

One of the most commonly used security criteria for daily operations is the $\mathrm{N}-1$ contingency (e.g., [62, 50]), where the system is able to continue operations without load shedding under any single component failure. A more stringent but less used criterion is $\mathrm{N}-k$ contingency, in which the system is required to sustain simultaneous failures of k electrically connected components [71]. While these security criterions effectively represent system operators' concerns for daily operations, they do not

capture the possible contingencies in extreme weather conditions or natural disaster occurrences, in which multiple components that are not electrically connected could fail simultaneously [112]. In this work, we generalize the $\mathrm{N}-k$ contingencies to any k simultaneous component failures and aim to prepare for the worst k failures. Finding defending strategy against the worst case among a set of adversary scenarios is often modeled as a bi-level interdiction framework, which is known as attacker-defender (AD) model (see, e.g., [93, 94, 79, 7]). In this context, the attackers are extreme weather events (natural disasters), trying to cause the most severe damages to system operations; the defenders are the re-dispatch actions that minimize the damages by redistributing power flows. However, the AD model helps to find near-optima but not the optimal protection plan against disruptive events, because it only seeks for the most critical set of assets, and hardening these critical assets is not necessarily the optimal protection plan [115, 24]. To obtain the optimal hardening decisions, the tri-level attacker-defender-attacker (DAD) model was initially proposed by [23]. The DAD model, which is an extension of the AD model, includes two interacting agents (attacker and defender) in three levels (see, e.g., [24, 16, 118, 51, 2, 117]). In the first level, the defender makes hardening decisions with a limited protection budget before disruptions. In the second level, the attacker disrupts the system to make the defender suffer from the highest cost or largest load shedding. In the third level, the defender aims to minimize the system cost against disruptions by taking corrective actions through re-dispaching the power output.

As renewable penetration continues to grow (e.g., it is predicted that 20% of nation's electricity is generated by wind energy by 2030 [73]), the power grids in the near future will have significant amount of renewable energy, creating challenges not only for operations but also for hardening planning. To capture the renewable energy uncertainty into optimization models, stochastic programming (see, e.g., [22, 104, 106, 91]) and robust optimization (see, e.g., [122, 128, 52]) approaches have

been extensively studied. However, stochastic programming and robust optimization approaches have shown disadvantages in practice. Stochastic programming approaches can be unreliable due to the blind assumption of probability distribution of the random parameter. In addition, stochastic programming becomes computationally challenging for a large number of the random parameter scenarios. Also, robust optimization approaches can be too conservative due to the consideration of the worst-case scenario of the uncertain parameter, which happens rarely.

One of the challenges for these methodologies to be practical is that, an accurate/complete knowledge on the probability distribution can hardly be obtained. Recently, distributionally robust and data-driven optimization (see, e.g., [37, 46, 55]) has been applied to power system operations. The advantages of distributionally robust optimization is that, it can handle uncertainties with partial information about the probability distribution. Distributionally robust optimization approaches have been successfully developed and implemented to solve power system optimization problems under uncertainty (see, e.g., [126, 111, 15, 121]). In this approach, by learning from a set of historical data or moment information of the random parameter, an ambiguity set for the unknown probability distribution of the random parameter is constructed. Then, the objective is to minimize the total cost under the worst-case distribution scenario within the constructed ambiguity set. Due to the consideration of the worst-case distribution, distributionally robust optimization leads to risk-averse and conservative solutions, as compared with the traditional stochastic approaches. However, the distributionally robust optimization in general is less conservative than the traditional robust approaches [46]. That is because, distributionally robust optimization takes advantage of data information to build an ambiguity set of distribution and considers the worst-case distribution in the ambiguity set to keep robustness, while robust optimization ignores the probability of random parameter scenarios, which usually leads to an unnecessarily high average cost. We will also numerically show this fact

in our case study.

In this study, we develop a defender-attacker-defender-based transmission system hardening planning (TSHP) model under both random disruptions (natural disasters) and uncertain wind power generation. Due to the availability of a considerable amount of historical data for wind power output to ISOs/RTOs, we deploy the distributionally robust optimization approach to formulate the wind output uncertainty. That is, we allow the ambiguity of the probability distribution of the wind output, and construct the ambiguity set for the unknown distribution. More specifically, to build the ambiguity set, we use the Wasserstein metric as a probability measure [85, 113], which also has many applications in transportation theory [87]. The built transmission system hardening planning model is a two-stage model which deals with the hardening decisions in the first stage and the re-dispatching decision in consideration with the worst-case disruption scenario and the worst-case wind output probability distribution in the second stage. The contributions of this research can be listed as follows:

1. A planning tool for power systems with high renewable generation capacity is developed. This tool, which considers both renewable generation uncertainty and multiple random disruptions, helps power system operators to efficiently allocate protective resources in order to reduce the vulnerability of the power transmission system against multiple transmission line contingencies caused by natural disasters or terrorist attacks, as well as maintain power system reliability with a large penetration of renewable energy.

2. A data-driven two-stage stochastic transmission system hardening planning model is developed. Using the Wasserstein metric, an ambiguity set is constructed for the unknown wind output distribution. It can efficiently utilize data information and reduce the conservativeness of the model. In addition, a tractable reformulation of the original formulation is proposed which can be

solved efficiently.

The remaining parts of this chapter are organized as follows: In Section 3.2, we define sets, parameters and variables. In Section 3.3, we describe how to utilize the historical data and construct the confidence set. Also, we develop a two-stage data-driven defender-attacker-defender model under both natural disaster and wind power uncertainties. In section 3.4, the proposed decomposition framework and the solution algorithm are presented. In section 3.5, numerical results are discussed. Finally, the research is concluded in section 3.6.

3.2 Nomenclature

A. Sets

$\mathcal{B}$	Index set of all buses.
$\mathcal{B}_i$	Index set of all buses directly connected to bus i.
$\mathcal{E}$	Index set of transmission lines.
$\mathcal{E}_A$	Index set of attacked transmission lines.
$\mathcal{E}_H$	Index set of hardened transmission lines.
$\mathcal{T}$	Index set of load blocks.

B. Parameters

L_{it}	Load shedding cost at bus i for load block t.
H_{ij}	Investment cost to harden transmission line (i, j).
U	The maximum number of lines affected by natural disasters simultaneously.

F_{ij} Flow capacity of transmission line (i, j).

C_i Generation unit capacity at bus i.

X_{ij} Reactance of transmission line (i, j).

θ_i^{min} Phase angle lower limit at bus i.

θ_i^{max} Phase angle upper limit at bus i.

d_{it} Demand at bus i for load block t.

w_{it} Renewable energy output at bus i for load block t.

C. Decision Variables

z_{ij} Binary decision variable to indicate whether transmission line (i, j) is hardened ($z_{ij} = 1$) or not ($z_{ij} = 0$).

u_{ij} Binary decision variable to indicate whether transmission line (i, j) is attacked ($u_{ij} = 0$) by natural disaster or not ($u_{ij} = 1$).

x_{it} Power generation at bus i for load block t.

$f_{ij,t}$ Power flow from bus i to bus j on transmission line (i, j) for load block t.

θ_{it} Phase angle at bus i for load block t.

s_{it} Load shedding at bus i for load block t.

3.3 Problem Formulation

In this research, we allow the probability distribution of wind output to be ambiguous because a particular distribution assumption of wind power output can be biased from

the actual one. We construct an ambiguity set for the true probability distribution of wind output which is centered at the reference distribution that is learned from a given historical data set.

3.3.1 Reference Distribution

We discuss the way to utilize the historical data and to obtain the reference distribution to estimate the true distribution. Given a set of historical data points, we let the histogram of data be the reference distribution of the wind power output. Without loss of generality, we assume the random wind power output $w(\xi)$ to be bounded above and below, within a supporting space Ω. We partition Ω into N bins $B_1, ..., B_N$, so that $\Omega = \bigcup_{n=1}^{N} B_n$. Given a set of historical data with size S, we obtain the reference distribution $\hat{P} = (\hat{p}^1, \hat{p}^2, ..., \hat{p}^N)$, where $\hat{p}^n = S_n/S$, $\forall n = 1, 2, \cdots, N$, and S_n denotes the frequency of data samples in bin n. It is worth noting that, in our data-driven modeling approach, the number of bins N is the only parameter that affects the computational complexity. Hence, as long as we do not change N, the computational complexity of our proposed data-driven approach remains unchanged.

3.3.2 Confidence Set Construction

We use the reference distribution to construct a confidence set (or ambiguity set) for the true distribution of wind output with confidence level β. Note that the reference distribution $\hat{P}$ is inherently different from the true distribution P. To measure the difference between P and $\hat{P}$, i.e., $d(P, \hat{P})$, we use the Wasserstein metric as the probability measure [85, 113]. Accordingly, $d(P, \hat{P})$ is defined as below:

$$d_{\mathrm{w}}(P, \hat{P}) := \inf_{\mathbb{Q}} \left\{ E_{\mathbb{Q}}[\mathrm{d}(w, \hat{w})] : P = \rho(w), \hat{P} = \rho(\hat{w}) \right\}, \tag{3.1}$$

where w and $\hat{w}$ are random wind power output associated with the true distribution and the reference distribution, respectively. $\mathrm{d}(w, \hat{w})$ is the distance between random

variables w and $\hat{w}$. $\mathbb{Q}$ denotes all joint distributions of w and $\hat{w}$ with marginal distributions P and $\hat{P}$. $\rho(.)$ denotes that P and $\hat{P}$ are distribution functions.

Then, we can construct a distribution-based ambiguity set using the Wasserstein metric as below:

$$\begin{aligned}\mathbb{D} &= \left\{ P \in \mathcal{M}_+ : d_{\mathrm{w}}(P, \hat{P}) \leq \varphi \right\} \\ &= \left\{ P \in \mathcal{M}_+ : \inf_{\mathbb{Q}} \left\{ E_{\mathbb{Q}}[\mathrm{d}(w, \hat{w})] : P = \rho(w), \hat{P} = \rho(\hat{w}) \right\} \leq \varphi \right\}, \end{aligned} \tag{3.2}$$

where $\mathcal{M}_+$ represents the set of all probability distributions and φ denotes the tolerance level of the distance, which is depending on the confidence level β and the size of the historical data S. The relationship between φ and S, under the Wasserstein metric, can be described by the following proposition (please see the Appendix for the proof):

Proposition 1 *Given a set of historical data of size S, N bins and a supporting space Ω with diameter D, the convergence rate between P and $\hat{P}$ under the Wasserstein metric is as follow:*

$$Pr(d_{\mathrm{w}}(P, \hat{P}) \leq \varphi) \geq 1 - 2N \exp(-4\varphi S/ND). \tag{3.3}$$

Accordingly, if the confidence level, i.e. the right-hand side of inequality (3.3), is set to be β, then we have

$$\varphi = \frac{ND}{4S} \log(\frac{2N}{1-\beta}). \tag{3.4}$$

Based on (3.4), as the size of historical data S increases, the value of φ decreases, i.e. the distance between the reference distribution and the true distribution $d(P, \hat{P})$ becomes smaller, and $\hat{P}$ converges to P.

We denote the central point of bin n as w^n, $n = 1, 2, \cdots, N$, which represent the discretized scenarios of the uncertain parameter. Based on the definition of Wasserstein metric and the construction of $d(P, \hat{P})$ in (3.2), we can reformulate the ambiguity

set $\mathbb{D}$ in (3.2) by the following constraints:

$$\sum_{n=1}^{N}\sum_{m=1}^{N} q^{nm}|w^m - w^n| \leq \varphi, \tag{3.5}$$

$$\sum_{n=1}^{N} q^{nm} = p^m, \forall m \tag{3.6}$$

$$\sum_{m=1}^{N} q^{nm} = \hat{p}^n, \forall n \tag{3.7}$$

$$\sum_{m=1}^{N} p^m = 1, \tag{3.8}$$

where p^n and $\hat{p}^n$ represent the true probability and reference probability of scenario n respectively, and q^{nm}, $n = 1, \cdots N, m = 1, \cdots N$ denotes the joint probability distribution (i.e., $\mathbb{Q}$ in (3.2)). Constraint (3.5) is a reformulation of (3.2). Constraints (3.6) and (3.7) represent that $p^m, m = 1, \cdots, N$ and $\hat{p}^n, n = 1, \cdots, N$ are marginal distributions of q^{nm}, $n = 1, \cdots N, m = 1, \cdots N$.

3.3.3 Data-Driven TSHP Framework

In this section, we develop a data-driven stochastic defender-attacker-defender model for the transmission system hardening planning problem considering uncertain wind power generation and unknown disruptive events, such as natural disasters.

As a matter of fact, although natural disasters (or terrorist attacks) happen infrequently and the related historical records are quite limited, they can bring catastrophic impacts when they happen. However, using historical records, we can claim that the number of transmission lines disrupted simultaneously is no more than U. Therefore, we formulate the uncertainty set of random disruptions as follows:

$$\mathbb{U} = \left\{ \sum_{(i,j)\in\mathcal{E}} (1 - u_{ij}) \leq U, u_{ij} \in \{0, 1\} \right\}. \tag{3.9}$$

In addition, natural disasters are intrinsically correlated in location. In order to consider this fact in our proposed model, we can restrict the random disruptions to

the area that is vulnerable to natural disasters by adjusting the uncertainty set $\mathbb{U}$. In this case, we replace $\mathcal{E}$ in (3.9) with $\mathcal{E}_v$, which denotes the transmission lines in the vulnerable area (see [117]).

To hedge against the risk brought by natural disasters and wind energy intermittency to the power system, we consider the joint worst case of disruption scenario in $\mathbb{U}$ and probability distribution of wind output in $\mathbb{D}$. Accordingly, we develop a data-driven two-stage stochastic model which considers the hardening decision variables in the first stage, and deals with power generation level, power flow, phase angle and load shedding variables by considering the worst-case transmission line disruption in $\mathbb{U}$ and the worst-case probability distribution of wind output in $\mathbb{D}$ in the second stage. We formulate a data-driven two-stage stochastic transmission system hardening planning model as follows:

$$\min_{z} \sum_{(i,j)\in\mathcal{E}} H_{ij} z_{ij} + \max_{u\in\mathbb{U}} \max_{P\in\mathbb{D}} E_P[\mathcal{Q}(z,u,w(\xi))] \tag{3.10}$$

$$s.t. \quad z_{ij} \in \{0,1\}, \forall (i,j) \in \mathcal{E}, \tag{3.11}$$

where,

$$\mathcal{Q}(z,u,w(\xi)) = \min_{p,f,\theta,s} \sum_{i} \sum_{t} L_{it} s_{it}(\xi) \tag{3.12}$$

$$s.t. \ (z_{ij} + u_{ij} - z_{ij}u_{ij})(\theta_{it}(\xi) - \theta_{jt}(\xi)) - X_{ij} f_{ij,t}(\xi) = 0,$$
$$\forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}, \tag{3.13}$$

$$-F_{ij}(z_{ij} + u_{ij} - z_{ij}u_{ij}) \le f_{ij,t}(\xi) \le F_{ij}(z_{ij} + u_{ij} - z_{ij}u_{ij}),$$
$$\forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}, \tag{3.14}$$

$$x_{it}(\xi) + \sum_{j\in\mathcal{B}_i(.,i)} f_{ji,t}(\xi) - \sum_{j\in\mathcal{B}_i(i,.)} f_{ij,t}(\xi) + s_{it}(\xi) = d_{it} - w_{it}(\xi),$$
$$\forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{3.15}$$

$$x_{it}(\xi) \le C_i, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{3.16}$$

$$\theta_i^{min} \le \theta_{it}(\xi) \le \theta_i^{max}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{3.17}$$

$$x_{it}(\xi), s_{it}(\xi) \ge 0, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{B}. \tag{3.18}$$

The objective function (3.10) is to minimize the system cost, i.e., hardening cost plus expected load shed cost under the worst-case disruption scenario and the worst-case distribution of wind output. Note here that in literature, hardening decisions or costs are often considered as budget constraints. In our paper, we include hardening costs in the objective function instead so that we can show different numbers of hardened lines led by different approaches. Constraints (3.13) represent the power flow in terms of phase angles. Constraints (3.14) are the transmission line flow capacity limits. In practice, hardening does not eliminate the chance of vulnerability but reduces it significantly to a very small level. Therefore, in this paper, we neglect the small chance of vulnerability of hardened lines and assume that the hardened lines are invulnerable to disruptions. Accordingly, the term $z_{ij} + u_{ij} - z_{ij}u_{ij}$ ensures that power flow constraints hold for any status of hardening and disruption. If line (i, j) is hardened ($z_{ij} = 1$), then $z_{ij} + u_{ij} - z_{ij}u_{ij} = 1$. So, this line is invulnerable and the power flow constraint holds. If line (i, j) is not hardened ($z_{ij,t} = 0$), then $z_{ij} + u_{ij} - z_{ij}u_{ij} = u_{ij}$. Hence, the power flow is depending on the attack scenario. If line (i, j) is attacked, i.e., $u_{ij} = 0$, then there is no flow on line (i, j) and (3.13) and (3.14) hold. If line (i, j) is not attacked, i.e., $u_{ij} = 1$, the power flow constraint also hold. Constraints (3.15) observe the power balance at each bus. Constraints (3.16) are the thermal generation capacity limits. Constraints (3.17) enforce the phase angle limit at each bus.

3.3.4 Linearizing Non-Linear Constraints

Note here that power flow constraints (3.13) and (3.14) are nonlinear. To linearize them, we fix the first-stage hardening decisions z_{ij} and consider them as input parameters in the second-stage problem. We consider two cases based on the hardening decisions derived from the first stage. In case one, we define set $\mathcal{E}_H = \{(i, j) \in \mathcal{E} | z_{ij} = 1\}$ as the subset of hardened lines. The transmission lines in $\mathcal{E}_H$ are not affected by

disruptions, and constraints (3.13) and (3.14) can be reformulated as:

$$(\theta_{it}(\xi) - \theta_{jt}(\xi)) - X_{ij} f_{ij,t}(\xi) = 0, \quad \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}_H, \tag{3.19}$$

$$-F_{ij} \leq f_{ij,t} \leq F_{ij}, \quad \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}_H. \tag{3.20}$$

In case two, for the transmission lines that are not in set $\mathcal{E}_H$, i.e, the lines are not protected, the power flows in these lines depend on the disruption status. In this case, we adopt the big-M method to linearize power flow constraints (3.13) and (3.14) as follows:

$$(\theta_{it}(\xi) - \theta_{jt}(\xi)) - X_{ij} f_{ij,t}(\xi) + M(1 - u_{ij}) \geq 0, \quad \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{3.21}$$

$$(\theta_{it}(\xi) - \theta_{jt}(\xi)) - X_{ij} f_{ij,t}(\xi) - M(1 - u_{ij}) \leq 0, \quad \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{3.22}$$

$$-F_{ij} u_{ij} \leq f_{ij,t} \leq F_{ij} u_{ij}, \quad \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{3.23}$$

Now, our data-driven model can be presented as the following program:

$$\begin{aligned} \min_{z} \quad & \sum_{(i,j) \in \mathcal{E}} H_{ij} z_{ij} + \max_{u \in \mathbb{U}, P \in \mathbb{D}} \sum_{m=1}^{N} p_{it}^{m} . \min_{p,f,\theta,s,o} \sum_{i} \sum_{t} L_{it} s_{it}(\xi^m) \\ s.t. \quad & \text{Constraints (3.11), (3.15)} - \text{(4.7)}. \end{aligned} \tag{3.24}$$

In the second-stage objective function, due to the independence of scenarios, we can interchange the second-stage minimization and summation (corresponding to the expectation term) operations. Then, we are able to reformulate our data-driven two-stage stochastic transmission system hardening planning problem as follows:

$$\begin{aligned} \min_{z} \quad & \sum_{(i,j) \in \mathcal{E}} H_{ij} z_{ij} + \max_{p,q,u} \min_{x,f,\theta,s,o} \sum_{m=1}^{N} \sum_{i} \sum_{t} p_{it}^{m} L_{it} s_{it}(\xi^m) \\ s.t. \quad & \text{Constraints (3.5)} - \text{(3.9)}, \\ & \text{Constraints (3.11), (3.15)} - \text{(4.7)}. \end{aligned} \tag{3.25}$$

3.4 Solution Methodology

To solve our proposed data-driven two-stage model, we employ a decomposition algorithm. We first reformulate our model into a tractable reformulation. Then, we

describe our proposed decomposition framework and the solution algorithm to solve the proposed data-driven TSHP model.

In order to reformulate the second-stage problem into a tractable formulation, we dualize the inner minimization problem and combine it to the outer maximization problem. Since the inner minimization is a linear program, which will lead to no dualization gap. The dual form of the inner minimization problem can be presented as follows:

$$\max_{\lambda,\gamma,\mu,\eta,\sigma,\tau,\upsilon} \sum_{m=1}^{N} \Bigg(\sum_{i}\sum_{t} \left((d_{it}^m - w_{it}^m)\lambda_{it}^m - C_i\gamma_{it}^m + \theta_i^{min}\bar{\eta}_{it}^m - \theta_i^{max}\tilde{\eta}_{it}^m\right) - \sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H} \sum_{t} \left(F_{ij}u_{ij}(\bar{\tau}_{ij,t}^m + \tilde{\tau}_{ij,t}^m) + M(1-u_{ij})(\bar{\sigma}_{ij,t}^m + \tilde{\sigma}_{ij,t}^m)\right) - \sum_{(i,j)\in\mathcal{E}_H} \sum_{t} \left(F_{ij}(\bar{\mu}_{ij,t}^m + \tilde{\mu}_{ij,t}^m)\right) \Bigg) \tag{3.26}$$

$s.t.$

$$\lambda_{it}^m - \gamma_{it}^m \leq 0, \forall t, \forall i, \quad \forall m \tag{3.27}$$

$$\lambda_{it}^m - \lambda_{jt}^m + \bar{\tau}_{ij,t}^m - \tilde{\tau}_{ij,t}^m - X_{ij}\bar{\sigma}_{ij,t}^m + X_{ij}\tilde{\sigma}_{ij,t}^m = 0, \quad \forall (i,j) \in \mathcal{E}\backslash\mathcal{E}_H, \forall t, \forall m \tag{3.28}$$

$$\lambda_{it}^m - \lambda_{jt}^m + \bar{\mu}_{ij,t}^m - \tilde{\mu}_{ij,t}^m - X_{ij}\upsilon_{ij,t}^m = 0, \quad \forall (i,j) \in \mathcal{E}_H, \forall t, \forall m \tag{3.29}$$

$$\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H(i,.)} (\bar{\sigma}_{ij,t}^m - \tilde{\sigma}_{ij,t}^m) + \sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H(.,i)} (\tilde{\sigma}_{ij,t}^m - \bar{\sigma}_{ij,t}^m) + \sum_{(i,j)\in\mathcal{E}_H(i,.)} \upsilon_{ij,t}^m - \sum_{(i,j)\in\mathcal{E}_H(.,i)} \upsilon_{ij,t}^m + \bar{\eta}_{it}^m - \tilde{\eta}_{it}^m = 0, \quad \forall t, \forall i, \forall m \tag{3.30}$$

$$\lambda_{it}^m - L_{it}p_{it}^m \leq 0, \quad \forall t, \forall i, \forall m \tag{3.31}$$

$$\bar{\eta}_{it}^m, \tilde{\eta}_{it}^m, \bar{\tau}_{it}^m, \tilde{\tau}_{it}^m, \bar{\mu}_{ij,t}^m, \tilde{\mu}_{ij,t}^m, \bar{\sigma}_{ij,t}^m, \tilde{\sigma}_{ij,t}^m \geq 0,$$

$$\gamma_{it}^m, p_{it}^m \geq 0, \lambda_{it}^m, \upsilon_{it}^m \text{ are free}, \quad \forall t, \forall i, \forall m \tag{3.32}$$

where λ_{it}^m and γ_{it}^m are dual variables of constraints (3.15) and (3.16), respectively; $\bar{\eta}_{it}^m$ and $\tilde{\eta}_{it}^m$ are dual variables of constraints (3.17); $\upsilon_{ij,t}^m$ are dual variables of constraints (4.2); $\bar{\mu}_{ij,t}^m$ and $\tilde{\mu}_{ij,t}^m$ are dual variables of constraints (4.6); $\bar{\sigma}_{ij,t}^m$ and $\tilde{\sigma}_{ij,t}^m$ are dual variables of constraints (4.4) and (4.5); $\bar{\tau}_{ij,t}^m$ and $\tilde{\tau}_{ij,t}^m$ are dual variables of constraints (4.7);

3.4.1 Linearizing the Objective Function

Here, we use big-M method to linearize the bilinear terms in the objective function (3.26). We let

$$-\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} u_{ij}\bar{\tau}^m_{ij,t} = -\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\bar{\nu}_{ij}, \tag{3.33}$$

$$-\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} u_{ij}\tilde{\tau}^m_{ij,t} = -\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\tilde{\nu}_{ij}. \tag{3.34}$$

Then, we have

$$-\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} F_{ij}u_{ij}(\bar{\tau}^m_{ij,t}+\tilde{\tau}^m_{ij,t}) = -\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H} F_{ij}(\bar{\nu}_{ij}+\tilde{\nu}_{ij}), \tag{3.35}$$

$$s.t.\ \ \bar{\nu}_{ij} \geq \sum_{m}\sum_{t}\bar{\tau}^m_{ij,t} - M(1-u_{ij}),\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.36}$$

$$\bar{\nu}_{ij} \geq -Mu_{ij},\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.37}$$

$$\tilde{\nu}_{ij} \geq \sum_{m}\sum_{t}\tilde{\tau}^m_{ij,t} - M(1-u_{ij}),\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.38}$$

$$\tilde{\nu}_{ij} \geq -Mu_{ij},\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H. \tag{3.39}$$

In addition, We let

$$\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} u_{ij}\bar{\sigma}^m_{ij,t} = \sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\bar{\epsilon}_{ij}, \tag{3.40}$$

$$\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} u_{ij}\tilde{\sigma}^m_{ij,t} = \sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\tilde{\epsilon}_{ij}. \tag{3.41}$$

Then, we have

$$\sum_{m}\sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H}\sum_{t} Mu_{ij}(\bar{\sigma}^m_{ij,t}+\tilde{\sigma}^m_{ij,t}) = \sum_{(i,j)\in\mathcal{E}\backslash\mathcal{E}_H} M(\bar{\epsilon}_{ij}+\tilde{\epsilon}_{ij}) \tag{3.42}$$

$$s.t.\ \ \bar{\epsilon}_{ij} \leq \sum_{m}\sum_{t}\bar{\sigma}^m_{ij,t} + M(1-u_{ij}),\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.43}$$

$$\bar{\epsilon}_{ij} \leq Mu_{ij},\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.44}$$

$$\tilde{\epsilon}_{ij} \leq \sum_{m}\sum_{t}\tilde{\sigma}^m_{ij,t} + M(1-u_{ij}),\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H, \tag{3.45}$$

$$\tilde{\epsilon}_{ij} \leq Mu_{ij},\ \ \forall(i,j)\in\mathcal{E}\backslash\mathcal{E}_H. \tag{3.46}$$

3.4.2 Decomposition Framework

We employ the Column-and-Constraint generation approach [120] in a decomposition framework, to solve the developed data-driven TSHP model. This approach is an iterative approach. In each iteration, the master problem (MP) is a relaxation of the original problem, which aims to find a lower bound of the problem and a current best hardening decision (may not be feasible). The sub-problem (SUB) is a reformulation of the second-stage problem, with the objective of obtaining an upper bound of the original problem and the worst-case scenarios. The iterative procedure continues until the difference between the upper bound and the lower bound is no more than a predefined level. As a result, the optimal hardening planning decision will be output and the worst-case wind output distribution and disruption scenario will be identified.

In our model, since the first-stage decisions are hardening decisions, they do not affect the feasibility of the second-stage problem. Moreover, since we allow load shedding in the model, the second-stage problem is always feasible for any hardening decision z derived from the first stage. Hence, there is no need for the first-stage decision feasibility check. Then, we have the following MP.

$$
\begin{aligned}
& \min_{z} \sum_{(i,j)\in\mathcal{E}} H_{ij} z_{ij} + \vartheta \\
s.t.\quad & \text{Constraints (3.11)}, \\
& \text{Constraints (3.15)} - \text{(4.10)}, \text{for each } \xi^m \\
& \text{Optimality cuts},
\end{aligned}
$$

where ϑ is set to be the second-stage objective value. Based on the reformulation of

(3.35) to (3.46), we can reformulate the sub-problem (SUB) as below:

$$
\begin{aligned}
\psi(z) = & \max_{\lambda,\gamma,\mu,\eta,\sigma,\tau,\epsilon,\upsilon,\nu,p,q,u} \sum_{m=1}^{N} \\
& \Bigg(\sum_{i} \sum_{t} \left((d_{it}^{m} - w_{it}^{m}) \lambda_{it}^{m} - C_i \gamma_{it}^{m} + \theta_i^{min} \bar{\eta}_{it}^{m} - \theta_i^{max} \tilde{\eta}_{it}^{m} \right) \\
& - \sum_{(i,j) \in \mathcal{E} \setminus \mathcal{E}_H} \sum_{t} \left(M(\bar{\sigma}_{ij,t}^{m} + \tilde{\sigma}_{ij,t}^{m}) \right) - \sum_{(i,j) \in \mathcal{E}_H} \sum_{t} F_{ij}(\bar{\mu}_{ij,t}^{m} + \tilde{\mu}_{ij,t}^{m}) \Bigg) \\
& + \sum_{(i,j) \in \mathcal{E} \setminus \mathcal{E}_H} \left(M(\bar{\epsilon}_{ij} + \tilde{\epsilon}_{ij}) - F_{ij}(\bar{\nu}_{ij} + \tilde{\nu}_{ij}) \right) \qquad (3.47) \\
& \text{Constraints } (3.5) - (3.9),\ (3.27) - (3.32), \\
& \text{Constraints } (3.36) - (3.39),\ (3.43) - (3.46).
\end{aligned}
$$

3.4.3 Column-and-Constraint Generation Algorithm

The Column-and-Constraint generation algorithm can be summarized in the following steps and in the flowchart presented in Fig. 3.1.:

1. Initialization. Set $k = 1, \vartheta = -\infty$.

2. Solve MP and get the first-stage optimal solution (the optimal hardening plan) z^*, ϑ and $\mathcal{E}_H$.

3. Solve SUB for the current hardening plan z^* to obtain the worst-case disruption scenario u^*, the worst-case wind output distribution p^* and $\psi(z^*)$.

4. If $\psi(z^*) \leq \vartheta$, stop. Current z^* is the optimal hardening plan. Output the result. Otherwise, set $k = k + 1$. For the disruption plan u^*, let subset $\mathcal{E}_A = \{(i,j) \in \mathcal{E} | u_{ij} = 0\}$ be the set of attacked transmission lines. Then, add the following

constraints to MP and go to step 2.

$$\vartheta \geq \sum_{m=1}^{N} \sum_{i} \sum_{t} p_{it}^{m*} L_{it} s_{it}(\xi^m), \tag{3.48}$$

$$(\theta_{it}(\xi^m) - \theta_{jt}(\xi^m)) - X_{ij} f_{ij,t}(\xi^m) + M(1 - z_{ij}) \geq 0,$$

$$\forall t, \forall m, \forall (i,j) \in \mathcal{E}_A, \tag{3.49}$$

$$(\theta_{it}(\xi^m) - \theta_{jt}(\xi^m)) - X_{ij} f_{ij,t}(\xi^m) - M(1 - z_{ij}) \leq 0,$$

$$\forall t, \forall m, \forall (i,j) \in \mathcal{E}_A, \tag{3.50}$$

$$(\theta_{it}(\xi^m) - \theta_{jt}(\xi^m)) - X_{ij} f_{ij,t}(\xi^m) = 0,$$

$$\forall t, \forall m, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_A, \tag{3.51}$$

$$-F_{ij} z_{ij} \leq f_{ij,t}(\xi^m) \leq F_{ij} z_{ij}, \quad \forall t, \forall m, \forall (i,j) \in \mathcal{E}_A, \tag{3.52}$$

$$-F_{ij} \leq f_{ij,t}(\xi^m) \leq F_{ij}, \quad \forall t, \forall m, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_A. \tag{3.53}$$

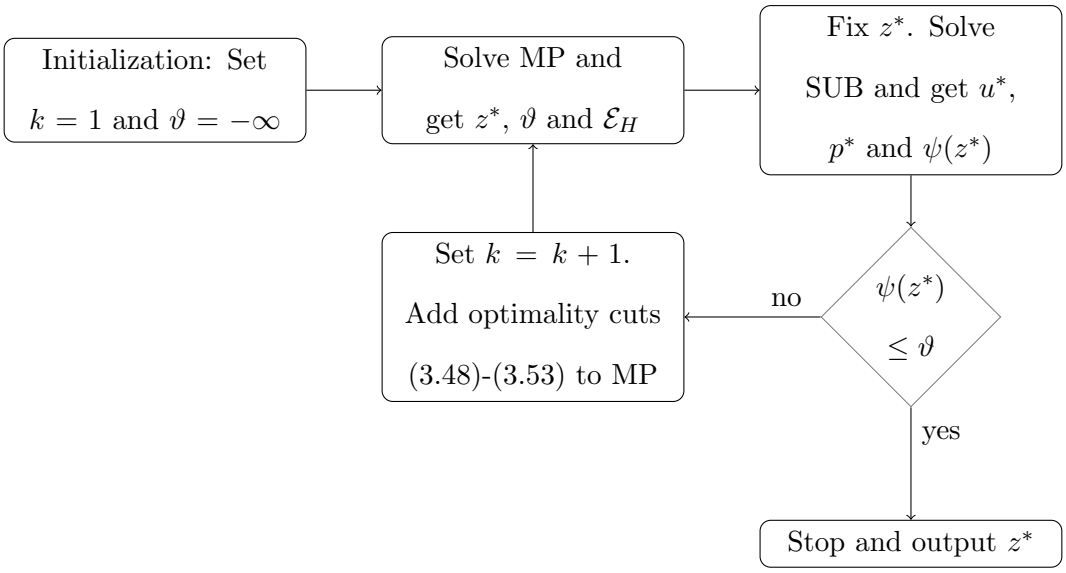

Figure 3.1: Flowchart of the Column-and-Constraint Generation algorithm

3.5 Case Study

In this section, we conduct numerical experiments to test the performance of the proposed approach. We apply our approach to a 24-node system [33], which is based

on IEEE one-area RTS-96 test system [49], and a modified IEEE 118-bus test system (available at http://motor.ece.iit.edu/data). According to [76], wind energy constitutes 5.6% of the total electricity generation across the united states in 2016. Accordingly, we modify both test systems by adding several wind power generation capacities, which account for 10% of the total generation capacity. We compare the system performance between the proposed data-driven transmission system hardening planning (DDTSHP) and Robust transmission system hardening planning (ROTSHP). We also discuss how the historical data and the ambiguity set can affect the conservativeness of the proposed data-driven approach. To implement the proposed algorithm, we use C++ and CPLEX 12.6 and run it on a computer with Intel(R) Xeon(R) 3.2 GHz and 8 GB memory.

3.5.1 Data Generation

In order to generate the set of historical data, we use the Monte Carlo simulation. For simulation convenience, we assume that wind outputs are independent for different load blocks. Also, for each load block, we assume that the unknown wind output follows a normal distribution with the forecasted wind output as the mean and 0.3 of the mean value as the standard deviation. To generate wind output scenarios, we generate samples for each wind farm and each time block and set the number of bins to be 5.

3.5.2 24-Bus System

This system, depicted in Fig. 3.2, consists of 24 buses, 12 generators, 34 transmission lines and 17 loads. We consider all generators to be thermal plants and add three wind farms at buses 10, 15 and 20. For convenience, we number transmission lines in Table 3.1.

Firstly, we compare the proposed DDTSHP with ROTSHP. We numerically show

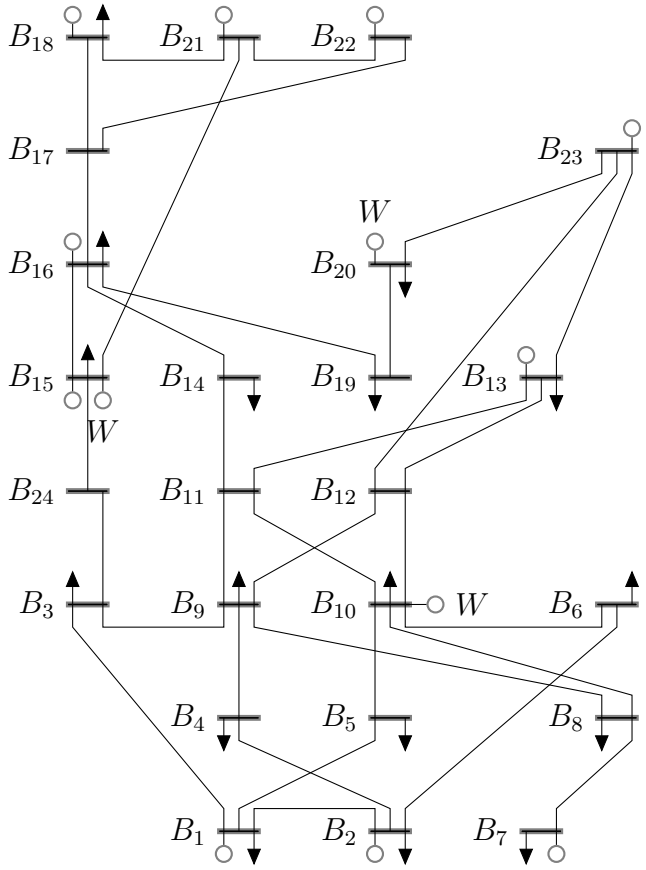

Figure 3.2: Modified 24-node system and the DDTSHP hardening plan for $U = 4$

DDTSHP obtains less conservative hardening plans compared with ROTSHP. To this end, we let the attack budget vary within a range of 1 to 10. We set the size of historical data set S and the confidence level β to be 100 and 0.99%, respectively. With a set of generated data in Section V-A, we solve the modified 24-node test system for different attack budgets and get the optimal first-stage hardening plans using both DDTSHP and ROTSHP. Then, we fix the optimal hardening decisions and solve the second-stage problem with a new set of randomly simulated wind output and attack scenarios for different attack budgets. In Table 3.2, we represent the optimal hardening plans obtained by DDTSHP and ROTSHP for various attack budgets. In Table 3.3, we report the number of hardened transmission lines (denoted by NHL), total cost (denoted by Obj) and the computational time in seconds (denoted by CPU)

Table 3.1: Transmission lines of 24-node system

No.	From	To	No.	From	To	No.	From	To
1	1	2	13	8	10	25	15	21
2	1	3	14	9	11	26	15	24
3	1	5	15	9	12	27	16	17
4	2	4	16	10	11	28	16	19
5	2	6	17	10	12	29	17	18
6	3	9	18	11	13	30	17	22
7	3	24	19	11	14	31	18	21
8	4	9	20	12	13	32	19	20
9	5	10	21	12	23	33	20	23
10	6	10	22	13	23	34	21	22
11	7	8	23	14	16			
12	8	9	24	15	16			

in Table 3.3. For example, the optimal hardening plan by DDTSHP (asterisked lines) for $U = 4$ is presented in Fig. 3.2. Notice that, from Table 3.3, there is no load shed for attack budget $U = 1$. However, as the attack budget U increases, the total cost increases for both DDTSHP and ROTSHP. We can also observe that DDTSHP results in less total cost compared with that of ROTSHP for all attack budgets. In addition, we see for most of attack budgets that DDTSHP's hardening plans include lesser number of hardened lines than those of ROTSHP. From these results, we can claim that the proposed DDTSHP is less conservative than ROTSHP. That is because DDTSHP utilizes the data information and considers the worst-case distribution of the wind output within the ambiguity set, while ROTSHP considers the worst-case wind output scenario. Hence, ROTSHP leads to a higher number of transmission lines to be hardened and therefore, this over-conservativeness results in higher hardening

cost and accordingly higher total cost.

Table 3.2: Hardening plans of DDTSHP versus ROTSHP for 24-node system

Attack Budget	Hardening Plan	
	DDTSHP	ROTSHP
2	(5, 23, 28)	(5, 23, 32, 33)
3	(6, 32, 33)	(11, 19, 32, 33)
4	(23, 32, 33)	(11, 18, 22, 32, 33)
5	(6, 11, 14, 18, 32, 33)	(11, 14, 18, 22, 32, 33)
6	(6, 11, 14, 18, 32, 33)	(6, 11, 14, 16, 18, 32, 33)
7	(6, 11, 14, 18, 32, 33)	(6, 11, 14, 16, 18, 32, 33)
8	(6, 11, 14, 18, 32, 33)	(6, 10, 11, 14, 17, 18, 21, 32, 33)
9	(6, 11, 14, 18, 32, 33)	(6, 10, 11, 14, 17, 18, 21, 32, 33)
10	(6, 11, 14, 18, 32, 33)	(6, 10, 11, 14, 17, 18, 21, 32, 33)

Secondly, we numerically illustrate and discuss how the size of historical data set can affect the conservativeness of the proposed DDTSHP. Accordingly, we conduct numerical experiments on the modified 24-node test system for different attack budgets from 2 to 10 and represent the results in Table 3.4. Here, we allow the size of historical data to vary between 50 to 10000 and set the confidence level β to be 99%. As shown in Table 3.4, as the size of historical data increases, the total cost decreases. That is because, according to equality (3.4), there is an inverse correlation between the value of φ and the size of historical data set S. As the number of historical observations increases, the value of φ decreases; and hence the confidence set $\mathbb{D}$ shrinks. From the theoretical point of view, eventually with an infinite number of historical observations, the value of φ goes to zero. In fact, in this case, the reference distribution converges to true distribution and the proposed data-driven approach becomes risk neutral. Moreover, we can see that larger attack budgets lead to higher

Table 3.3: DDTSHP versus ROTSHP for 24-node system

Attack	DDTSHP			ROTSHP		
Budget	NHL	Obj ($m)	CPU(s)	NHL	Obj ($m)	CPU(s)
1	0	0	0.8	0	0	0.8
2	3	2.6209	8.5	4	2.8865	5.5
3	3	2.6465	18.8	4	3.4887	10.9
4	3	3.1073	33.1	5	4.4717	15.4
5	6	5.0757	30.6	6	5.3680	13.8
6	6	5.4582	27.6	7	6.0558	15.3
7	6	5.5317	28.1	7	6.4745	10.8
8	6	5.9142	21.0	9	6.6830	10.3
9	6	6.0783	24.2	9	6.6925	12.1
10	6	6.1923	23.0	9	6.7198	11.6

total costs.

Thirdly, we assess the effect of the ambiguity set $\mathbb{D}$ on the performance of the proposed data-driven approach. Here, we set the size of historical data S to be 100 and test our DDTSHP on the modified 24-node test system. Note here that, according to (3.4), as the value of β increases, the value of φ increases. Thus, a larger value of β leads to a bigger ambiguity set $\mathbb{D}$. Therefore, we can control the size of ambiguity set by adjusting the confidence level β. We let the confidence level β to vary within a range from 0.5 to 0.99 and represent the associated system cost for different values of attack budget in Table 3.5. From the results we can see. as the value of β increases, our proposed data-driven approach gets more conservative and the total system cost increases (See Table 3.5).

Table 3.4: Effects of the historical data on total cost ($m)

Attack	# of data					
Budget	50	100	500	1000	5000	10000
2	2.6256	2.6209	2.6126	2.6114	2.6105	2.6104
3	2.6480	2.6465	2.6448	2.6445	2.6443	2.6442
4	3.4665	3.4600	3.4486	3.4470	3.4458	3.4457
5	5.7041	5.6910	5.6808	5.6794	5.6782	5.6780
6	5.9675	5.9536	5.9444	5.9430	5.9417	5.9415
7	6.0234	6.0137	6.0031	6.0012	5.9998	5.9996
8	6.2500	6.2393	6.2292	6.2276	6.2265	6.2350
9	6.5043	6.4928	6.4825	6.4810	6.4798	6.4796
10	6.5745	6.5628	6.5533	6.5523	6.5507	6.5505

Table 3.5: Effects of the ambiguity set on total cost ($m)

Attack	Confidence level β					
Budget	0.5	0.7	0.8	0.9	0.95	0.99
2	2.6170	2.6173	2.6175	2.6182	2.6188	2.6209
3	2.6453	2.6456	2.6457	2.6458	2.6459	2.6464
4	3.4506	3.4525	3.4547	3.4553	3.4552	3.4595
5	5.6841	5.6852	5.6861	5.6871	5.6880	5.6919
6	5.9474	5.9476	5.9481	5.9499	5.9508	5.9541
7	6.0045	6.0057	6.0070	6.0079	6.0107	6.0122
8	6.2312	6.2325	6.2328	6.2343	6.2354	6.2482
9	6.4847	6.4867	6.4879	6.4880	6.4889	6.4928
10	6.5557	6.5567	6.5574	6.5593	6.5613	6.5675

3.5.3 118-Bus System

In this section, we illustrate the effectiveness of our proposed DDTSHP through conducting experiments on a larger system. The 118-bus test system consists of 118 buses, 33 generators and 186 transmission lines. We consider all generators to be thermal plants and add five wind farms at buses 20, 40, 60, 80 and 100. We consider a range of 2 to 10 for the attack budget. We also let the size of historical data S be 100 and the confidence level β be 99%. Then, we follow the same simulation procedure as the one in the previous subsection and compare the performance of DDTSHP with ROTSHP. We report the results in Table 3.6, which has the same information as Table 3.3. We observe that as the attack budget increases, the total system cost increases too. Moreover, compared to ROTSHP, we see that DDTSHP leads to lesser number of hardened lines and lower total cost, for most of the attack budgets. These results come from the fact that DDTSHP is less conservative than ROTSHP thanks to the data information utilization. These observations and results admit the results from the experiments on 24-node system.

3.6 Summary

In this study, we developed an approach to deal with the stochastic transmission system hardening planning problem in the presence of wind generation uncertainty and multiple simultaneous disruptive events. Motivated by the considerable amount of historical data available to power system operators, we proposed a data-driven approach which learns from the wind output historical data and employs the Wasserstein metric to construct an ambiguity set for the unknown wind output probability distribution. Our data-driven two-stage model can obtain a robust hardening decisions by considering the joint worst-case disruptive scenario and wind output distribution. A decomposition framework based on Column-and-Constraints generation method is used to solve the proposed model. We showed through the numerical experiments that

Table 3.6: DDTSHP versus ROTSHP for 118-bus system

Attack	DDTSHP			ROTSHP		
Budget	NHL	Obj ($m)	CPU(s)	NHL	Obj ($m)	CPU(s)
2	4	1.8855	24.3	4	1.8855	13.8
3	4	1.9014	79.2	4	2.2706	30.2
4	4	2.2177	95.3	7	3.0744	42.8
5	6	3.0061	226.2	8	3.7661	78.7
6	6	3.2099	291.5	8	3.8470	78.4
7	6	3.4583	272.0	8	4.0783	115.9
8	6	3.6620	442.9	8	4.2694	105.0
9	6	3.9104	489.9	8	4.7832	90.6
10	10	5.1314	529.1	12	5.5279	98.7

although our approach is risk-averse, it leads to less conservative hardening decisions than the robust optimization approach. In addition, we showed as the conservativeness of our approach depends on the number of available historical data and the confidence level we prefer.

CHAPTER 4

RELIABILITY ANALYSIS OF TRANSMISSION SYSTEM HARDENING VIA DATA-DRIVEN OPTIMIZATION

Increasing complexity of power transmission networks has led power systems to be more vulnerable to cascading failures. Thus, hardening and reliability assessment of such complex networks have become a must. In addition, the commonly used $\mathrm{N}-1$ security criterion does not guarantee the system reliability against possible cascading failures. In this study, given a hardening plan, we develop two models to evaluate the reliability of the power transmission system under $\mathrm{N}-k$ security criterion. In the first model, we quantify the probability of no load shedding in the system to assess the possibility of load curtailment. Then, to have a better insight of the amount of load shed, in the second model, we use the conditional value-at-risk (CVaR) as a risk measure to evaluate the system reliability. To do a reliability assessment, the information of contingency probabilities are required. However, such probability information is usually unknown and cannot be estimated precisely. Therefore, in this study, we assume the probability of contingencies unknown and ambiguous. Then, we construct an ambiguity set for the unknown probability distribution of contingencies. Our approaches are robust because they analyze the transmission system reliability with respect to the worst-case distribution in the ambiguity set. We formulate both models as bi-level programs and solve them by Bender's decomposition technique. Finally, we conduct numerical experiments on 6-bus and IEEE 118-bus test systems to show the effectiveness of our proposed approaches.

4.1 Problem Description and Literature Review

Rapid growth of electric power systems and creation of very complex interconnected power transmission networks have made power systems more vulnerable to cascading failures and large blackouts than in the past [77]. Large blackouts are among the most catastrophic disasters that threaten the US economy through massive economic damage of tens of billions of dollars yearly [63]. Transmission system outages (mostly caused by severe weather conditions [10, 64], aging [61, 65] and terrorist attacks [118, 16]) are among the major causes of large blackouts [6]. As a matter of fact, in the deregulated electricity market, the transmission system is utilized such that it operates near its limits, i.e., as economically as possible [35], in which case an initial line outage may affect other system components and result in cascading failures and large blackouts (for example, blackouts in February 2008 in Florida [72] and September 2011 in North America [69]). Therefore, due to the criticality of the electric power industry to the national economy and society in general, hardening planning and reliability evaluation of power transmission systems is of significant importance.

According to North American Electric Reliability Council (NERC), reliability is the degree of power system performance under which customers' electricity demand is supplied and delivered under the accepted standards [95]. This definition of reliability contains two concepts: 1) Adequacy: the ability of a power system to supply the customers' electricity demand via available generation units and transmission systems and reserves. 2) Security: the ability of a power system to keep working after some contingencies such as transmission line outages or equipment failures. In the power system literature, the most popular reliability indices are loss of load probability (LOLP), loss of load expectation (LOLE) and expected energy not supplied (EENS) [17]. Moreover, two frequently used techniques for power systems reliability analysis are Monte Carlo simulations and contingency analysis.

In Monte Carlo simulation techniques, system reliability assessment is carried out by sampling system component states. Component state samples are generated randomly either from an estimated distribution of failures (using historical data of component failures) i.e., random sampling (e.g., [19]), or by considering component states transition probabilities i.e., sequential sampling (e.g., [18]). Moreover, to improve the efficiency of Monte Carlo simulation algorithms, several variations of this technique have recently been proposed, such as variance reduction techniques [20, 130], least square support vector classifier [86], artificial neural network [34], fuzzy Monte Carlo technique [28], cross-entropy methods [48], Latin hypercube sampling [98], etc.

In contingency analysis, a set of contingency states is used to examine the power system reliability. The contingency set is created by taking $\mathrm{N}-1, ..., \mathrm{N}-k, k = 1, 2, 3, ...$ security criteria into account, where k denotes the number of simultaneously component (generation units, transmission lines and etc.) outages. In industry practice, $\mathrm{N}-1$ is the most widely used security criterion by most power systems around the world [112, 62]. However, it just guarantees the normal operations of the system under only a single component failure. Although the probability of two or more simultaneous component failures is very small, it may lead to very severe cascading failures and blackouts if the simultaneous failures happen. Therefore, to establish more reliable system operations against multiple simultaneous contingencies, revised NERC reliability standards [70] require system operators to apply $\mathrm{N}-k, k \geq 2$ security criterion in their analysis. However, for $k \geq 2$, the combinatorial nature of contingency states makes the full contingency enumeration almost impossible for even medium size systems and moderate values of k. For more discussions on the complexity of full contingency enumeration, readers are referred to [5].

To mitigate the computational burden of $\mathrm{N}-k$ contingency analysis, contingency selection procedures have been proposed. Briefly, contingency selection is the process of identifying the critical components and constructing a contingency list that

includes very serious single and multiple contingencies. Some prior studies (see, e.g., [77, 31, 30, 54]) estimate the probability of system component failures using the historical data of component failures. Then, contingencies with higher probabilities form the contingency list. Reference [77] presents a statistical method to estimate the probability of transmission line failures and to identify vulnerable lines in a transmission system. Reference [31] proposes a method based on substation configuration and probability analysis of protection system failures to form a $N-k$ contingency list. Reference [30] uses three probabilitic models to estimate the probabilities of multiple transmission line contingencies. Also, [54] develops a data mining based method for contingency analysis. However, some other works (see, e.g., [60, 35, 114]) consider the consequence (e.g. load shedding) of contingency scenarios to form the contingency list. Reference [60] presents a fast and reliable heuristic algorithm based on iterative pruning to identify critical $N-2$ contingencies. Reference [35] develops two contingency screening algorithms to determine the threatening $N-2$ contingencies without solving the full contingency set. Reference [114] develops a method that uses a small number of representative constraints instead of enumerating exponentially many constraints for $N-k$ contingency analysis. In addition, many optimization-based approaches have also been utilized for contingency selection procedures and protective resources allocation, i.e., hardening planning. For instance, [93, 94, 7, 79], among others, develop bi-level programs to identify the most critical transmission system components under $N-k$ contingency criterion. Also, [118, 16, 2], among others, develop tri-level optimization models to simultaneously identify the most critical system components and the optimal hardening strategy to mitigate (enhance) the transmission system vulnerability (reliability) against intentional attacks. Moreover, other approaches such as a random chemistry algorithm [40] and a combined neural network and evolutionary algorithm [53] have recently been developed to study $N-k$ contingency analysis.

All above-mentioned reliability analysis approaches either estimate probabilities of component failures or assume same probability to each component and deal with the consequence of failures. However, historical data for equipment failures are usually rare and the estimated equipment failure rates, which rely on the expert information, are subject to errors. Even with sufficient historical data, accurate estimation of failure rate is extremely difficult. On the contrary, in this work, we allow the probability of contingencies to be unknown and ambiguous. We apply the distributionally robust optimization concept (see, e.g., [37, 46, 55]) for reliability analysis of power transmission systems. Distributionally robust optimization approaches have recently received attention from power system specialists for various problems such as unit commitment [126, 127], transmission expansion planing [9] and reserve scheduling [110, 111], etc. In this approach, instead of fixing the probabilities of component failures, we allow them to vary within a set of probability distributions, which is the ambiguity set. We develop two distributionally robust optimization models to determine the system reliability, for a given hardening plan, based on two reliability indices. In the first model, we develop a model to quantify the worst-case probability under which both system security and system adequacy are satisfied (we call it worst-case no-load-shed probability). In the second model, we use the conditional value-at-risk (CVaR) as its risk measure. CVaR, which is a well-known risk measure, has gained considerable attention in finance and insurance industries [88]. We propose a model to evaluate the worst-case CVaR associated with the worst-case contingency distribution in the ambiguity set. The contributions of this study can be listed as below:

1. We propose two reliability analysis models for transmission system hardening plans that hedge against the uncertainty (inaccuracy) associated with the estimation of probabilities of component failures. Instead of relying on estimates, our models consider the worst-case reliability with the worst-case failure distribution in the ambiguity set. Moreover, the conservatism of the proposed models

can be adjusted by system operators.

2. Our models are formulated in such a way that decomposition techniques can be easily employed to solve them by commercial solvers. Here, we apply Bender's decomposition technique to solve both models.

3. We conduct expensive numerical experiments on a modified 6-bus and IEEE 118-bus test systems. We compare both models for various contingency levels and different hardening plans to test the performance of our models.

The remaining parts of this chapter are organized as follows: In Section 4.2, we define sets, parameters and variables. In Section 4.3, we first discuss how to construct the ambiguity set for the unknown contingency distribution. Then, we develop a distributionally robust optimization model to quantify the worst-case no-load-shed probability of the system. Afterward, we formulate a distributionally robust optimization model to evaluate the transmission system reliability based on the worst-case CVaR risk measure. In Section 4.4, we describe the proposed decomposition framework and the solution algorithms. In section 4.5, numerical results are presented and discussed. Finally, we conclude this study in section 4.6.

4.2 Nomenclature

A. Sets

$\mathcal{B}$	Index set of all buses.
$\mathcal{B}_i$	Index set of all buses directly connected to bus i.
$\mathcal{B}_i(.,i)$	Index set of all incoming transmission lines to bus i.
$\mathcal{B}_i(i,j)$	Index set of all outgoing transmission lines from bus i.
$\mathcal{E}$	Index set of transmission lines.

$\mathcal{E}_H$ Index set of hardened transmission lines.

$\mathcal{T}$ Index set of load blocks.

B. Parameters

F_{ij} Flow capacity of transmission line (i, j).

C_i Generation capacity at bus i.

X_{ij} Reactance of transmission line (i, j).

θ_i^{min} Phase angle lower limit at bus i.

θ_i^{max} Phase angle upper limit at bus i.

d_{it} Demand at bus i for load block t.

C. Decision Variables

x_{it} Power generation at bus i for load block t.

$f_{ij,t}$ Power flow from bus i to bus j on transmission line (i, j) for load block t.

θ_{it} Phase angle at bus i for load block t.

s_{it} Load shedding at bus i for load block t.

D. Random Variables

v_{ij} Binary variable indicating whether transmission line (i, j) is under contingency ($v_{ij} = 0$) or not ($v_{ij} = 1$).

4.3 Problem Formulation

In this research, we aim to develop risk assessment models for the transmission system hardening decisions under distributional uncertainty of N$-k$ contingencies. Basically, we consider the following constraints to analyze the reliability of a power transmission system:

$$x_{it} + \sum_{j \in \mathcal{B}_i(.,i)} f_{ji,t} - \sum_{j \in \mathcal{B}_i(i,.)} f_{ij,t} + s_{it} = d_{it}, \ \forall i, \forall t, \tag{4.1}$$

$$(\theta_{it} - \theta_{jt}) - X_{ij} f_{ij,t} \geq 0, \ \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}_H, \tag{4.2}$$

$$(\theta_{it} - \theta_{jt}) - X_{ij} f_{ij,t} \leq 0, \ \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}_H, \tag{4.3}$$

$$(\theta_{it} - \theta_{jt}) - X_{ij} f_{ij,t} + M(1 - v_{ij}) \geq 0, \ \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{4.4}$$

$$(\theta_{it} - \theta_{jt}) - X_{ij} f_{ij,t} - M(1 - v_{ij}) \leq 0, \ \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{4.5}$$

$$-F_{ij} \leq f_{ij,t} \leq F_{ij}, \ \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E}_H, \tag{4.6}$$

$$-F_{ij} v_{ij} \leq f_{ij,t} \leq F_{ij} v_{ij}, \forall t \in \mathcal{T}, \forall (i,j) \in \mathcal{E} \backslash \mathcal{E}_H, \tag{4.7}$$

$$x_{it} \leq C_i, \ \forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{4.8}$$

$$\theta_i^{min} \leq \theta_{it} \leq \theta_i^{max}, \ \forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{4.9}$$

$$x_{it}, s_{it} \geq 0, \ \forall t \in \mathcal{T}, \forall i \in \mathcal{B}, \tag{4.10}$$

where, constraints (4.1) observe the power supply adequacy at each bus. Constraints (4.2) to (4.3) represent the relationship between DC power flow and phase angle for hardened lines. Constraints (4.4) to (4.5) represent the relationship between DC power flow and phase angle for unhardened lines. That is, if the unhardened lines are under contingency, i.e., $v_{ij} = 0$, the relationship of power flow and phase angle as suggested in (4.2) to (4.3) does not necessarily hold. Similarly, constraints (4.6) and (4.7) observe power flow capacities for hardened and not hardened lines, respectively. Constraints (4.8) and (4.9) impose power generation capacities and phase angle limits, respectively.

4.3.1 Ambiguity Set

Due to the fact that there are exponentially many $\mathrm{N}-k$ contingency scenarios and enumerating all of them is almost impossible (it takes exponential time), in this study, we create a list of N contingencies. To this end, two groups of contingencies can be considered: 1) the most probable contingencies and 2) the highest impact contingencies (contingencies with the worst consequences). Then, based on the system operator's preferences, either of these or a mixture of them can be used to create the contingency list. The most probable contingencies and their probabilities can be identified by utilizing existing methods in the literature (see, e.g., [77, 31, 30, 54] among others). The high impact contingency scenarios can also be identified by optimization-based methods (see, e.g., [118, 93, 94, 7, 79], among others). These contingency scenarios and their assigned probabilities form a discrete distribution $\hat{P}$, which we call the reference distribution. Considering the fact that the true probability of $\mathrm{N}-k$ contingencies is unknown, using the reference distribution $\hat{P}$, we construct an ambiguity set for the ambiguous probability distribution of $\mathrm{N}-k$ contingencies. We define the following ambiguity set $\mathbb{D}$:

$$\mathbb{D} := \big\{ P \in \mathcal{M}_+ : d(P, \hat{P}) \leq \varphi, \tag{4.11}$$

$$\underline{P} \leq P \leq \overline{P} \big\}, \tag{4.12}$$

where $\mathcal{M}_+$ represents the set of all probability distributions. In (4.11), $d(P, \hat{P})$ denotes the probability distance between any arbitrary distribution $P \in \mathbb{D}$ and the reference distribution $\hat{P}$. Also, φ denotes the tolerance level for the probability distance. To measure $d(P, \hat{P})$, several probability metrics, such as L_1, L_∞, Wasserstein metric, etc., can be utilized. For more detail, interested readers are referred to [125]. In this study, we use L_1 norm. According to Proposition 1 in [126], given a set of historical data, the following equation defines the relationship between the size of

data and the value of φ under L_1 norm:

$$\varphi = \frac{N}{2S} \log \frac{2(N)}{1-\beta} \ , \tag{4.13}$$

where S and β denote the size of historical data and the confidence level, respectively. However, if in practice, the data information for contingencies is very limited and we do not have enough data to learn a practical value of φ, φ can also be decided by the system operators based on their judgment or preference on the conservativeness level. Since the worst-case probability distribution is considered in $\mathbb{D}$ in our proposed models, to control the conservativeness, we limit the probability of each contingency occurring with inequalities (4.12), where $\underline{P}$ and $\overline{P}$ are the lower and the upper probability limit matrices. Then, the ambiguity set $\mathbb{D}$ under L_1 norm can be represented as below:

$$\mathbb{D} := \big\{ P \in \mathbb{R}_+^N : \sum_{n=1}^{N} |p_n - \hat{p}_n| \leq \varphi, \tag{4.14}$$

$$\underline{p}_n \leq p_n \leq \overline{p}_n, \forall n \big\}, \tag{4.15}$$

where p_n, $\underline{p}_n$ and $\overline{p}_n$ denote the unknown probability, the lower probability limit and the upper probability limit for contingency scenario n, respectively. By setting $\underline{p}_n = \hat{p}_n - \delta$ and $\overline{p}_n = \hat{p}_n + \delta$, we can adjust the conservativeness of our models by changing parameter δ. Note here this ambiguity set also implicitly contains the correlation information of the contingencies, since it is constructed based on the historical data, which captures the correlation of contingencies automatically. Fig. 4.1 illustrates one sample of the ambiguity set $\mathbb{D}$ with five contingency scenarios. The black solid lines show the reference probabilities of contingency scenarios, and the dash lines show the ambiguity of the true probabilities from the reference ones.

4.3.2 Worst-Case No-Load-Shed Probability (WNLP)

In this section, given a hardening plan, we intend to quantify the worst-case probability of a no load shedding occurrence over the entire system. For notational brevity,

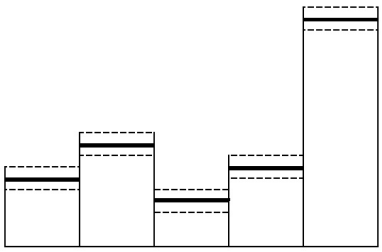

Figure 4.1: An example of the ambiguity set

we let

$$s_{i,t}(x,f) = x_{it} + \sum_{j \in \mathcal{B}_i(.,i)} f_{ji,t} - \sum_{j \in \mathcal{B}_i(i,.)} f_{ij,t} - d_{it},\ \forall i, \forall t. \tag{4.16}$$

Then, to satisfy the adequacy requirement, we need to determine decision variables x_{it}, $f_{ij,t}$ and θ_{it} such that the probability of no-load-shed over the entire network is maximized. So, we consider the following program:

$$\max_{x,f,\theta \in \mathbf{X}} \ Pr\big(s_{i,t}(x,f) \geq 0, \forall i, \forall t\big), \tag{4.17}$$

where $\mathbf{X}$ denotes the set of all feasible solutions satisfying security requirements, i.e. system constraints (4.2) to (4.10). In addition, notice that we have a joint probabilistic objective function, in (4.17), to quantify the no-load-shed probability for the whole system. We can reformulate the objective function in (4.17) as

$$\max_{x,f,\theta \in \mathbf{X}} \ Pr\big(\min_{i,t} s_{i,t}(x,f) \geq 0\big). \tag{4.18}$$

The probabilistic objective function above can be expressed as:

$$\max_{x,f,\theta \in \mathbf{X}} \ E_P\Big[\mathbf{1}_{[0,\infty)}\big(\min_{i,t} s_{i,t}(x,f)\big)\Big], \tag{4.19}$$

where indicator function $\mathbf{1}_{[0,\infty)}(x)$ equals 1 if $x \in [0,\infty)$ and zero otherwise. Note that, in this study, instead of making an assumption (by estimation or expert information) for contingency scenarios distribution P and taking expectation over P, we let P be unknown and ambiguous, but belong to the ambiguity set $\mathbb{D}$. Then, the strategy here is to allow P to act adversely against the maximization of the expected value in

(4.19), i.e., bringing robustness into the model. Hereby, we compute the worst-case probability of no-load-shed (WNLP) by the following distributionally robust bi-level min-max problem:

$$\min_{P\in\mathbb{D}} \max_{x,f,\theta\in\mathbf{X}} \; E_P\Big[\mathbf{1}_{[0,\infty)}\big(\min_{i,t} s_{i,t}(x,f)\big)\Big]. \tag{4.20}$$

By defining new variable $y_n = \mathbf{1}_{[0,\infty)}\big(\min_{i,t} s^n_{i,t}(x,f)\big)$ for each contingency scenario n and using big-M method, we can obtain the following equivalent reformulation of (4.20):

$$\min_{P\in\mathbb{D}} \max_{x,f,\theta\in\mathbf{X}} \sum_{n=1}^{N} p_n y_n, \tag{4.21}$$

$$s^n_{i,t}(x,f) \geq -M(1-y_n), \forall i, \forall t, \forall n, \tag{4.22}$$

$$y_n \in \{0,1\}, \forall n. \tag{4.23}$$

4.3.3 Worst-Case Conditional Value-at-Risk (WCVaR)

In this section, given a hardening plan, we propose a new reliability analysis scheme for power transmission systems by considering the conditional value-at-risk (CVaR) risk measure, which is also known as Mean Excess Loss, Mean Shortfall, or Tail VaR [88]. CVaR is closely related to the popular measure of risk VaR (an upper percentile of the loss distribution) and is defined as the weighted average of VaR and losses strictly exceeding VaR. CVaR, in comparison with VaR, has nice mathematical properties such as translational invariance, sub-additivity, convexity and homogeneity, which make this risk measure coherent and practical to use in optimization problems. Moreover, from the above definition, VaR can never be more than CVaR. For more details, readers are referred to [88]. We let

$$L_{i,t}(x,f) = \big(-s_{i,t}(x,f)\big)^+, \forall i, \forall t \tag{4.24}$$

be the loss (load shedding) associated with decision variables x_{it}, $f_{ij,t}$ and θ_{it}, where $(x)^+ = \max\{0,x\}$. Then, since we aim to evaluate the whole transmission system

reliability, we define a joint loss function as $L(x,f) = \sum_i \sum_t L_{i,t}(x,f)$. Therefore, γ–CVaR for the joint loss function $L(x,f)$ and for the probability level $\gamma \in (0,1)$ can be presented as [88]:

$$\gamma\text{–CVaR} = \min_{\alpha} \alpha + \frac{1}{1-\gamma} E_P \Big[\big(L(x,f) - \alpha \big)^+ \Big], \tag{4.25}$$

where P denotes the probability distribution of random contingencies. To ensure adequacy, we need to determine decision variables x_{it}, $f_{ij,t}$ and θ_{it} such that γ–CVaR is minimized. Accordingly, we develop the following program:

$$\min_{x,f,\theta \in \mathbf{X},\alpha} \gamma\text{–CVaR}, \tag{4.26}$$

where $\mathbf{X}$ denotes the same set as that in section 4.3.2. Following the procedure in section 4.3.2, we propose a distributionally robust optimization approach to minimize the worst-case γ–CVaR (γ-WCVaR), based on the worst-case probability distribution of random contingencies in the ambiguity set $\mathbb{D}$. Accordingly, we have the following bi-level max-min program:

$$\max_{P \in \mathbb{D}} \min_{x,f,\theta \in \mathbf{X},\alpha} \alpha + \frac{1}{1-\gamma} E_P \Big[\big(L(x,f) - \alpha \big)^+ \Big]. \tag{4.27}$$

Here, notice that there exists two $(.)^+$ terms in the objective function of γ-WCVaR model (4.27), i.e., $\big(L(x,f) - \alpha\big)^+$ and $L(x,f)$, according to (4.24). To linearize $L_{i,t}(x,f)$, we use auxiliary variables z_n for each contingency scenario n and add the following constraints to γ-WCVaR model (4.27):

$$z_{it}^n \geq -s_{i,t}^n(x,f), \forall i, \forall t, \forall n, \tag{4.28}$$

$$z_{it}^n \geq 0, \forall i, \forall t, \forall n. \tag{4.29}$$

Also, to linearize $\big(L(x,f) - \alpha\big)^+$ in (4.27), we use auxiliary variables u_n for each contingency scenario n and add the following constraints to γ-WCVaR problem (4.27):

$$u_n \geq \sum_i \sum_t z_{it}^n - \alpha, \forall n, \tag{4.30}$$

$$u_n \geq 0, \forall i, \forall t, \forall n. \tag{4.31}$$

We eventually attain the following reformulation of γ-WCVaR problem (4.27):

$$\max_{P\in\mathbb{D}} \min_{x,f,\theta\in\mathbf{X},\alpha} \alpha + \frac{1}{1-\gamma}\sum_{n=1}^{N} p_n u_n, \tag{4.32}$$

$$s.t.\ \text{Constraints (4.28) to (4.31).} \tag{4.33}$$

4.4 Solution Methodology

In this section, we describe our solution approaches for solving both proposed distributionally robust optimization models: WNLP (i.e., (4.21) to (4.23)) and β-WCVaR (i.e., (4.32) and (4.33)). We employ Bender's decomposition algorithm to solve both models. We explain the solution algorithms in detail in the following subsections.

4.4.1 Solution Approach for WNLP Model

Based on WNLP problem formulation, (4.21) to (4.23), we utilize the following Bender's decomposition framework to solve this problem. We define the master problem and the subproblem as follows:

Master Problem

We first represent the linear reformulation of (4.14) in the ambiguity set $\mathbb{D}$, (4.14) and (4.15), as the following inequalities:

$$\sum_{n=1}^{N} k_n \leq \varphi, \tag{4.34}$$

$$k_n \geq p_n - \hat{p}_n, \forall n, \tag{4.35}$$

$$k_n \geq \hat{p}_n - p_n, \forall n, \tag{4.36}$$

where variables k_n represent $|p_n - \hat{p}_n|$. The master problem for WNLP model is represented as:

$$\text{(WNLP-MP)} \quad \min_{p} \ \vartheta_1, \tag{4.37}$$

$$s.t. \ \sum_{n=1}^{N} p_n = 1, \tag{4.38}$$

$$\text{Constraints (4.15) and (4.34) to (4.36)}, \tag{4.39}$$

$$\text{Cutting planes}, \tag{4.40}$$

where ϑ_1 denotes the objective value of the subproblem that we will discuss in the next subsection. Also, constraint (4.38) ensures that variables p_n represent scenario probabilities.

Subproblem

In each iteration of the algorithm, given the solution p^* obtained in the master problem, we solve subproblem $\psi_1(p^*)$ as below:

$$\text{(WNLP-SP)} \quad \psi_1(p^*) = \max_{x,f,\theta} \ \sum_{n=1}^{N} p_n^* y_n \tag{4.41}$$

$$s.t. \ \text{Constraints (4.22) - (4.23)}, \tag{4.42}$$

$$x, f, \theta \in \mathbf{X}. \tag{4.43}$$

Solution Algorithm

Since we allow load shedding in the system, the subproblem is always feasible. In addition, the master problem solution p^* only appears in the subproblem objective function and does not affect the subproblem feasibility. So, we do not need to check the solution feasibility of the master problem. Therefore, to generate optimality cuts, in each iteration, for fixed WNLP-MP solutions p^* and ϑ_1^*, WNLP-SP is solved to obtain $\psi_1(p^*)$. Then, for the case $\psi_1(p^*) > \vartheta_1^*$, the following optimality cut is added

to WNLP-MP:

$$\vartheta_1 \geq \sum_{n=1}^{N} p_n y_n^* \tag{4.44}$$

The solution algorithm is summarized in the flowchart shown in Fig.4.2.

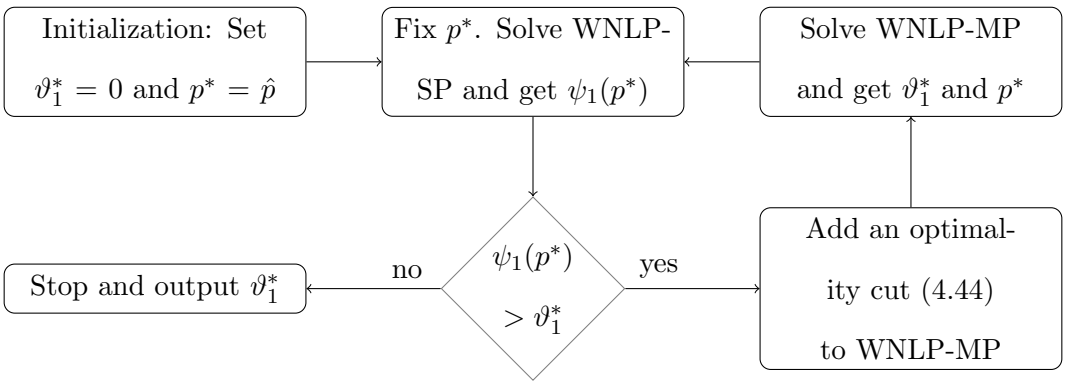

Figure 4.2: Solution algorithm for WNLP model

4.4.2 Solution Approach for WCVaR Model

For WCVaR model, the problem formulation (i.e., (4.32) to (4.33)) has a similar structure to WNLP model. Hence, a similar solution procedure based on Bender's decomposition is applied to solve the proposed model.

Master Problem

For WCVaR model, we can represent the master problem (WCVaR-MP) as below:

$$\text{(WCVaR-MP)} \quad \max_{p} \ \vartheta_2, \tag{4.45}$$

$$s.t. \text{ Constraints (4.38) to (4.40)}, \tag{4.46}$$

where ϑ_2 denotes the objective value of the subproblem presented in the next subsection.

Subproblem

For the solution p^* obtained by solving WCVaR-MP, we have the following subproblem $\psi_2(p^*)$:

$$\text{(WCVaR-SP)} \quad \psi_2(p^*) = \min \ \alpha + \frac{1}{1-\gamma} \sum_{n=1}^{N} p_n^* u_n \tag{4.47}$$

$$s.t. \ \text{Constraints (4.28) to (4.31)}, \tag{4.48}$$

$$x, f, \theta \in \mathbf{X}. \tag{4.49}$$

Solution Algorithm

Similarly, the feasibility of the master problem solution p^* to WCVaR-SP is always guaranteed. So, only optimality cuts are needed. In each iteration, for the fixed WCVaR-MP solutions p^* and ϑ_2^*, we obtain $\psi_2(p^*)$. Then, we check whether $\psi_2(p^*) < \vartheta_2^*$. If so, we add an optimality cut to WCVaR-MP as below:

$$\vartheta_2 \leq \alpha^* + \frac{1}{1-\gamma} \sum_{n=1}^{N} p_n u_n^* \tag{4.50}$$

We can illustrate the solution algorithm for WCVaR model by the flowchart in Fig.4.3.

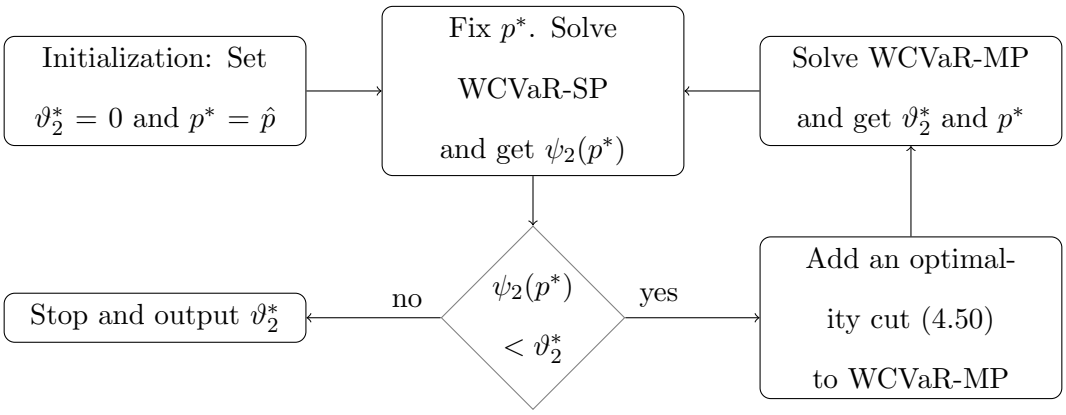

Figure 4.3: Solution algorithm for WCVaR model

4.5 Case Study

In this section, to show the effectiveness and efficiency of our proposed approaches, we present numerical experiments on a modified 6-bus test system and IEEE 300-bus test system (available at http://www.maths.ed.ac.uk/optenergy/LocalOpt/300busnetwork_other.html). We use C++ and CPLEX 12.6 on a computer with Intel(R) Xeon(R) 3.2 GHz and 8 GB memory to implement all the experiments.

For computational simplicity, we set the probability of each contingency to be 0.01, and we consider the critical contingencies with high impact in the contingency list to increase the robustness of our approaches. Therefore, our contingency list includes the worst-consequence contingency scenarios and the scenario of no line outage occurrence. We use the robust optimization approach proposed in [118] to identify the worst-consequence contingency scenarios.

4.5.1 6-Bus System

We test a modified 6-bus system consisting of three generation units, seven transmission lines and six load nodes. The 6-bus system specifications are presented in Tables 4.1 and 4.2.

Table 4.1: Bus data for 6-bus system

Bus No.	Generation Capacity (MW)	Load (MW)
1	270	100
2	0	100
3	200	100
4	0	100
5	300	100
6	0	100

Table 4.2: Line data for 6-bus system

Line No.	From Bus	To Bus	Reactance (pu)	Line Flow (MW)
1	1	2	0.037	200
2	1	4	0.016	200
3	2	3	0.1015	175
4	2	4	0.117	175
5	3	6	0.0355	175
6	4	5	0.037	200
7	5	6	0.127	200

We test this system for various contingency levels. We construct three contingency lists for $k = 1, 2, 3$. For $k = 1$, the contingency list includes all single transmission line contingency scenarios and a no-contingency scenario. For $k = 2$, we consider a contingency list of 11 scenarios (10 critical contingency states out of all contingencies with $k = 1, 2$ and a no-contingency state). Also, for $k = 3$, we construct a contingency list of 16 scenarios (15 critical contingency states out of all contingencies with $k = 1, 2, 3$ and a no-contingency state). We report the worst-case no-load-shed probability (WNLP) and the worst-case CVaR (WCVaR) (for $\beta = 0.95$) in Tables 4.3 and 4.4, respectively. Here, we also use several hardening plans with different budgets (i.e. the number of hardened lines) to show our model's capability for examining the performance of different hardening decisions. For this test system, we observe that the system reliability is guaranteed under $\mathrm{N}-1$ security criterion under both WNLP and WCVaR reliability indices. But, this is not the case for $k = 2, 3$. Nevertheless, by putting hardening plans into action, we are able to bring higher levels of reliability into the system. In Tables 4.3 and 4.4, the first column represents hardening decisions, where NH denotes no transmission line has been hardened. For instance, plan (4,6)

indicates that transmission lines 4 and 6 are hardened. However, different hardening plans lead to different levels of reliability. For example, plan (4,6), compared with plan (2,4) in N − 2 case, or plan (4,5,6), compared with plan (2,4,5) in N − 3 case, show higher WNLP but less WCVaR (better reliability indices).

Table 4.3: WNLP for 6-bus system

Hardening Plan	Security Criterion		
	N − 1	N − 2	N − 3
NH	1.000	0.9910	0.9905
(2)	1.000	0.9920	0.9915
(6)	1.000	0.9930	0.9926
(2,4)	1.000	0.9940	0.9936
(4,6)	1.000	1.0000	0.9948
(2,4,5)	1.000	1.0000	0.9949
(4,5,6)	1.000	1.0000	1.0000

Table 4.4: WCVaR for 6-bus system

Hardening Plan	Security Criterion		
	N − 1	N − 2	N − 3
NH	0.00	447.12	542.64
(2)	0.00	378.48	470.64
(6)	0.00	268.08	327.84
(2,4)	0.00	314.14	327.36
(4,6)	0.00	0.00	67.68
(2,4,5)	0.00	0.00	73.44
(4,5,6)	0.00	0.00	0.00

4.5.2 300-Bus System

In this section, we test our reliability analysis models by conducting experiments on a larger test system, i.e. IEEE 300-bus test system, which consists of 300 buses, 69 generators and 411 transmission lines. We test our models based on $N - 3$ transmission line contingencies. Accordingly, we create a contingency list of 46 scenarios, including 45 worst-consequence contingency scenarios among all contingencies with $k = 1, 2, 3$ and a no-line-outage scenario. In our experiments, we evaluate our model's performance by using different hardening plans with various hardening budgets.

Sensitivity analysis of φ

We conduct experiments for different values of φ (the distribution distance tolerance level) to see how the value of φ affects our model's performance. Here, we set the value of δ (the scenario probability tolerance level) to be 0.005. Tables 4.5 and 4.6 represent the values of WNLP and WCVaR (for $\beta = 0.95$), respectively. In Table 4.5, we observe that as the number of hardened lines increases (more hardening budget), the worst-case probability of no-load-shed increases. From Table 4.6, we see that with more hardened lines, the worst-case CVaR value decreases. In other words, with more hardening budget, the transmission system reliability increases. We also observe that as the value of φ decreases, the value of WNLP increases (WCVaR decreases). In fact, with smaller values of φ the ambiguity set of probability distributions gets tighter and both WNLP and WCVaR models become less conservative. Therefore, the objective value of WNLP increases (WCVaR decreases). In addition, by comparing the results in Tables 4.5 and 4.6, another important observation is that although the values of WNLP for plans (208) and (316) are the same, there is a significant difference between the value of WCVaR for these plans. This observation admits that the worst-case CVaR can bring a better insight of the system reliability than WNLP. The computational times (denoted by T), for all settings, are also presented

in Tables 4.5 and 4.6.

Table 4.5: Effects of φ on WNLP

Hardening Plan	$\varphi = 0.001$		$\varphi = 0.005$		$\varphi = 0.01$	
	WNLP	T(s)	WNLP	T(s)	WNLP	T(s)
NH	0.8330	446	0.8310	473	0.8285	457
(316)	0.8430	507	0.8410	498	0.8385	510
(208)	0.8430	514	0.8410	468	0.8385	494
(208, 316)	0.8540	592	0.8520	511	0.8495	503
(208, 316, 118)	0.8651	518	0.8631	521	0.8606	530
(208, 316, 118, 311)	0.8762	555	0.8742	547	0.8717	542
(208, 316, 118, 311, 342)	0.8874	605	0.8854	567	0.8829	586

Table 4.6: Effects of φ on WCVaR

Hardening Plan	$\varphi = 0.001$		$\varphi = 0.005$		$\varphi = 0.01$	
	WCVaR	T(s)	WCVaR	T(s)	WCVaR	T(s)
NH	13152	682	14542	913	16279	662
(316)	11169	644	12562	963	14305	674
(208)	6352	667	7557	997	9045	1081
(208, 316)	4294	655	5360	918	6676	1045
(208, 316, 118)	4035	685	5116	932	6452	1016
(208, 316, 118, 311)	3724	626	4837	848	6196	1028
(208, 316, 118, 311, 342)	3245	573	4378	815	5775	954

Sensitivity analysis of δ

Here, we assess the effects of the value of δ on the performance of our models. We use the same contingency list and the same hardening plans as section 4.5.2. We set the value of φ to be 0.01. We report WNLP and WCVaR values for different values of δ, in Tables 4.7 and 4.8, respectively. We see that as the value of δ increases the value of WNLP decreases (WCVaR increases). That is because, as the value of δ increases the ambiguity set becomes larger and our models get more conservative. Notice again, infrequency of component failures and the scarcity of failure historical data lead to relatively large values of φ and over conservatism. In such cases, as shown in Tables 4.7 and 4.8, our model's conservatism can be adjusted by changing the value of δ.

Table 4.7: Effects of δ on WNLP

Hardening	$\delta = 0.0005$		$\delta = 0.001$		$\delta = 0.005$	
Plan	WNLP	T(s)	WNLP	T(s)	WNLP	T(s)
NH	0.8330	451	0.8325	481	0.8285	457
(316)	0.8425	473	0.8415	505	0.8385	510
(208)	0.8425	480	0.8415	490	0.8385	494
(208, 316)	0.8525	488	0.8505	486	0.8495	503
(208, 316, 118)	0.8625	528	0.8606	522	0.8606	530
(208, 316, 118, 311)	0.8725	602	0.8717	529	0.8717	542
(208, 316, 118, 311, 342)	0.8829	575	0.8829	582	0.8829	586

Comparison with simulation method

In this subsection, we compare our proposed approaches with the traditional simulation method (denoted by Sim). We numerically show that the proposed WNLP and WCVaR approaches are more reliable than Sim. First, we generate 1000 sample con-

Table 4.8: Effects of δ on WCVaR

Hardening Plan	$\delta = 0.0005$		$\delta = 0.001$		$\delta = 0.005$	
	WCVaR	T(s)	WCVaR	T(s)	WCVaR	T(s)
NH	15624	671	15799	947	16279	662
(316)	13035	645	13575	929	14305	674
(208)	7924	1065	8477	1022	9045	1081
(208, 316)	5505	1014	6061	902	6676	1045
(208, 316, 118)	5275	1035	5836	917	6452	1016
(208, 316, 118, 311)	5018	1018	5580	878	6196	1028
(208, 316, 118, 311, 342)	4482	956	5178	892	5775	954

tingency scenarios as the historical record of contingencies, and obtain the reference distribution $\hat{P}$ and the ambiguity set $\mathbb{D}$ (see section 4.3.1) based on the historical data. Then, we use Sim to get the optimal generation level, power flow and phase angle $(x^*, f^*, \theta^*)_{\text{Sim}}$ with the reference distribution $\hat{P}$; solve WNLP to get the optimal generation level, power flow and phase angle $(x^*, f^*, \theta^*)_{\text{WNLP}}$ and worst-case distribution P-WNLP; and solve WCVaR to get the optimal generation level, power flow and phase angle $(x^*, f^*, \theta^*)_{\text{WCVaR}}$ and worst-case distribution P-WNLP. Notice, to get $(x^*, f^*, \theta^*)_{\text{Sim}}$, we use constraints (4.1) to (4.10), for each scenario n, with the objective of $\min_{x,f,\theta} \sum_{n=1}^{\mathcal{N}} \hat{p}_n s_n$. We then fix the decisions (x^*, f^*, θ^*) generated by Sim, WNLP, WCVaR accordingly and test them with 100000 new generated contingency scenarios. We compare the load shedding amount using the obtained (x^*, f^*, θ^*) for Sim, WNLP, WCVaR and report the results in Table 4.9 for varies of hardening plans. We observe that the operations decisions obtained by both WNLP and WCVaR lead to less amount of load shedding compared to the one by Sim. That is because the

decisions obtained by WNLP and WCVaR are based on the worst-case probability distribution of contingencies, which are more robust and reliable than the traditional simulation approach.

Then we test the case that the operations decisions (x, f, θ) are adjustable in real-time operations. In this case, we compare the amount of necessary adjustments in (x, f, θ) to minimize the load shedding for three approaches respectively. We report the results in Table 4.10 for different hardening plans. We observe that both WNLP and WCVaR's operations decisions require less adjustments than the ones of Sim, which means the decisions of WNLP and WCVaR are more robust than the ones obtained by Sim and therefore lead to less real-time operational costs.

Table 4.9: Load shedding (MW)

Hardening Plan	Sim	WNLP	WCVaR
NH	2136.9	1373.27	1421.4
(208)	1444.4	1027.0	729.0
(208, 316)	972.8	555.4	493.2
(208, 316, 118)	896.9	479.6	417.4
(208, 316, 118, 311)	801.0	423.8	377.2
(208, 316, 118, 311, 342)	735.7	376.8	358.9

In addition, we show the proposed method is computationally efficient than the traditional simulation approach. First, we set $\beta = 0.95$ and get the worst-case no-load-shed probabilities via WNLP model. Then, we generate 100 samples of 1000 contingency scenarios and obtain the no-load-shed probability for each sample, and then obtain the 95% lower-tailed confidence interval of no-load-shed probability by Sim. We report the bounds of the 95% confidence interval and their corresponding CPU times in Table 4.11 for different hardening plans. We see that the bounds of the confidence intervals (i.e., worst-case no-load-shed probabilities) are very close to

Table 4.10: Amount of adjustment

Hardening Plan	Sim	WNLP	WCVaR
NH	1322.8	559.2	607.4
(208)	1013.4	596.0	297.9
(208, 316)	672.0	264.6	192.4
(208, 316, 118)	635.2	217.8	155.6
(208, 316, 118, 311)	574.9	197.7	151.1
(208, 316, 118, 311, 342)	547.1	188.3	170.3

the ones of WNLP with $\beta = 0.95$ (i.e., the case of ambiguity set with 95% confidence level). However, we observe that the CPU times of Sim are significantly higher than those of WNLP. These results show that, by using WNLP model, we are able to obtain a lower bound for no-load-shed probability without conducting a huge number of experiments and simulations and it is more computationally efficient.

Table 4.11: 95% confidence interval of no-load-shed probability using Sim

Hardening Plan	Sim		WNLP	
	CI	T(s)	$\varphi = 0.005$	T(s)
NH	(0.8307, ∞)	>3600	0.8324	458
(208)	(0.8415, ∞)	>3600	0.8415	481
(208, 316)	(0.8530, ∞)	>3600	0.8525	487
(208, 316, 118)	(0.8635, ∞)	>3600	0.8645	504
(208, 316, 118, 311)	(0.8745, ∞)	>3600	0.8744	543
(208, 316, 118, 311, 342)	(0.8854, ∞)	>3600	0.8846	568

4.6 Summary

In this study, we proposed two reliability analysis schemes for the power transmission system hardening under distributional uncertainty of random contingencies and under $\mathrm{N}-k$ security criterion. We used the distributionally robust optimization concept to develop two optimization models to assess the transmission network reliability for a given transmission hardening plan. First, we developed a model to quantify the worst-case probability of no load shedding occurrence over the entire system. In the second model, we utilized the conditional value-at-risk (CVaR) as a risk measure and computed the worst-case CVaR to evaluate the transmission system reliability. In both models, instead of assuming a fixed probability estimate (by the historical data or expert information) for contingency scenarios, we let the ambiguous probability distribution of contingencies belong to an ambiguity set. Then, we considered the worst-case probability distribution in the ambiguity set to make a robust reliability analysis of the system. We formulated both models in bi-level programs and employed Bender's decomposition algorithm to solve them. By numerical experiments, we showed that our models are capable of distinguishing more effective hardening plans from others. Also, we observed that as the size of historical data increases, the conservatism of our models decreases. In addition, we showed that, in case of scarcity of historical data, the conservativeness of our proposed models can be adjusted by the system operators.

CHAPTER 5

POWER SYSTEM SCHEDULING VIA DATA-DRIVEN OPTIMIZATION

Rapid integration of cheap, clean but highly intermittent wind energy into power systems brings challenges to ISOs to maintain the system reliability. Stochastic programs may result in biased and unreliable unit commitment (UC) and economic dispatch (ED) decisions by fixing the probability distribution of wind output. Robust optimization approaches sacrifice system's cost-effectiveness in exchange of reliable UC and ED schedules. In this chapter, we develop a data-driven chance-constrained two-stage stochastic UC model to bridge the gap between stochastic programming and robust optimization. Without any particular assumption of wind output distribution, the data-driven chance constraint limits the worst-case chance of load imbalance to be no more than a specified tolerance, by taking advantage of historical data. We apply Column-and-Constraints Generation to solve our model. By experiments, we show the effectiveness of our model and the value of data.

5.1 Problem Description and Literature Review

The US government long-term financial incentives, such as tax credit programs and retirement plans, to promote investments in clean and cheap renewable energy (e.g. wind power) and to replace pollutant thermal plants with clean ones, have led to rapid renewable energy installments. Due to these stimulus plans, it is predicted that wind power will reach 20% of total energy generation across the country in 2030 [73]. However, unpredictable and intermittent nature of renewable energy affects the power

system stability and reliability. In this regard, ISOs have been traditionally utilizing excessive online reserves to prepare for the case that the actual wind power output is much lower than its predicted level. References [38] and [80] investigate the amount of required reserve for power systems with large amount of installed wind capacity. They show that the more wind capacity is integrated with the system, the higher reserve capacity should be scheduled. With the increasing penetration of wind energy into the grid, purely increasing the ancillary service deployment is not practical.

Recently, stochastic programming approaches have been widely applied to address the uncertainties in UC and ED. In [12], a stochastic linear programming model addressing security-constrained unit commitment with large amount of wind power capacity is presented. References [102] and [103] are extensions of [12] considering uncertain demand and equipment outages. In addition, two-stage stochastic models are commonly used to address wind power uncertainty, which typically consider UC in the first stage before the wind power output is known, and ED in the second stage after the wind power output is realized [107, 91, 83]. Furthermore, in order to reduce load shedding and renewable energy curtailment, risk-averse stochastic unit commitment models integrating chance constraints and expected value constraints have been successfully developed [81, 108, 129]. However, for stochastic programming, the probability distribution of the unknown parameters are usually assumed known; in practice, the distribution information is usually unknown and the UC decisions obtained with the inaccurate distribution assumption may affect the system reliability and cost efficiency.

In addition, robust optimization approaches have been successfully developed for UC and ED under uncertainty [57, 56, 128]. In robust optimization, the optimal schedules are obtained by considering the worst-case scenario of the unknown parameter which varies within an uncertainty set. In fact, this approach acquires system robustness by sacrificing the cost effectiveness of the system. Moreover, robust opti-

mization does not utilize the historical data to a large extent because uncertainty set construction needs limited information about the random parameter.

To address the reliability issues of stochastic programming and over conservatism of robust optimization, data-driven and distributionally robust optimization approaches have recently been developed [37, 46, 55]. In these approaches, the unknown probability distribution of the random parameter is allowed to vary in a confidence set, which is constructed by learning from a given historical data set. Though data-driven approaches are still risk-averse approaches since they consider the worst-case distribution in the confidence set, the conservatism is generally less than that in robust optimization. Moreover, as the size of data increases, the conservatism of this approach decreases. These approaches have recently been applied to several operational problems under uncertainty in power systems [126, 15, 111].

In order to take advantage of considerable amount of wind output historical data available for ISOs, in this chapter, we propose a data-driven chance-constrained stochastic UC (DDCHC) under wind uncertainty. We formulate it as a two-stage model, in the first stage consists of the traditional stochastic unit commitment problem and data-driven chance constraint that is used to restrict the probability of power imbalance, and in the second stage, the penalty cost due to energy imbalance is considered for the case the actual power output differs from the one committed. To conclude the contributions, (1) We develop a data-driven chance-constrained two-stage stochastic model under wind uncertainty which utilizes available historical data to obtain reliable but cost efficient generation schedules. (2) We show by experiments that our approach solutions are more reliable than the traditional chance-constrained UC (CCUC) solutions. Also, although our approach is risk-averse, its conservatism decreases as the size of data increases.

The remaining parts of this chapter are organized as follows: In Section 5.2, we define sets, parameters and variables. In section 5.3, we describe the mathematical

formulation of the model, the confidence set construction, and the chance constraint reformulation. In section 5.4, we discuss the solution methodology and efficient algorithm to solve the problem. In section 5.5, we show our numerical experiments and results to verify the effectiveness of our model. Finally, in section 5.6, we conclude this study.

5.2 Nomenclature

A. Parameters

SU_i^b	Start-up cost of generation unit i at bus b.
SD_i^b	Shut-down cost of generation unit i at bus b.
MU_i^b	Minimum up-time of generation unit i at bus b.
MD_i^b	Minimum down-time of generation unit i at bus b.
$\underline{C}_i^b$	Lower limit of generation capacity for generation unit i at bus b.
$\overline{C}_i^b$	Upper limit of generation capacity for generation unit i at bus b.
$\overline{R}_i^b$	Ramp-up rate limit of generation unit i at bus b.
$\underline{R}_i^b$	Ramp-down rate limit of generation unit i at bus b.
F_{ij}	Flow capacity of the transmission line (i, j) connecting bus i and bus j.
FD_{ij}^b	Flow distribution factor for the transmission line connecting bus i and bus j, based on the net injection at bus b.
d_t^b	Demand at bus b in time t.
$w_t^b(\xi)$	Wind power output at bus b in time t for scenario ξ.

π Load imbalance tolerance in the chance constraint.

ϵ Risk level of energy imbalance.

$F_i(.)$ Fuel cost of generation unit i.

L_t Penalty cost per unit of energy imbalance in time t.

B. Variables

y_{it}^b Binary variable indicating whether generation unit i is on $(= 1)$ or off $(= 0)$ in time t.

u_{it}^b Binary variable indicating whether generation unit i is started up $(= 1)$ or not $(= 0)$ in time t.

v_{it}^b Binary variable indicating whether generation unit i is shut down $(= 1)$ or not $(= 0)$ in time t.

x_{it}^b Amount of electricity generated by generation unit i at bus b in time t.

$s_t(\xi)$ Amount of energy imbalance (shortage or oversupply) in time t corresponding to scenario ξ.

5.3 Problem Formulation

5.3.1 Chance-Constrained Two-Stage Formulation

The above-mentioned chance-constrained two-stage UC can be formulated as follows:

$$\min \sum_t \sum_b \sum_i (SU_i^b u_{it}^b + SD_i^b v_{it}^b + F_i(x_{it}^b)) + E[Q(y,u,v,g,\xi)] \tag{5.1}$$

$$s.t. \quad -y_{i(t-1)}^b + y_{it}^b - y_{ik}^b \le 0, \quad \forall t, \forall b, \forall i, \forall k : 1 \le k-(t-1) \le MU_i^b \tag{5.2}$$

$$y_{i(t-1)}^b - y_{it}^b + y_{ik}^b \le 1, \quad \forall t, \forall b, \forall i, \forall k : 1 \le k-(t-1) \le MD_i^b \tag{5.3}$$

$$-y_{i(t-1)}^b + y_{it}^b - u_{it}^b \le 0, \quad \forall t, \forall b, \forall i, \tag{5.4}$$

$$y_{i(t-1)}^b - y_{it}^b - v_{it}^b \le 0, \quad \forall t, \forall b, \forall i, \tag{5.5}$$

$$\underline{C}_i^b y_{it}^b \le x_{it}^b \le \overline{C}_i^b y_{it}^b, \quad \forall t, \forall b, \forall i, \tag{5.6}$$

$$x_{it}^b - x_{i(t-1)}^b \le (2 - y_{i(t-1)}^b - y_{it}^b)\underline{C}_i^b + (1 + y_{i(t-1)}^b - y_{it}^b)\overline{R}_i^b, \quad \forall t, \forall b, \forall i, \tag{5.7}$$

$$x_{i(t-1)}^b - x_{it}^b \le (2 - y_{i(t-1)}^b - y_{it}^b)\underline{C}_i^b + (1 - y_{i(t-1)}^b + y_{it}^b)\underline{R}_i^b, \quad \forall t, \forall b, \forall i, \tag{5.8}$$

$$-F_{ij} \le \sum_b FD_{ij}^b (w_t^b(\xi) + \sum_r x_{rt}^b - d_t^b) \le F_{ij}, \quad \forall t, \forall (i,j), \tag{5.9}$$

$$Pr(-\pi \le \sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b \le \pi) \ge 1 - \epsilon, \quad \forall t, \tag{5.10}$$

$$y_{it}^b, u_{it}^b, v_{it}^b \in \{0,1\}, x_{it}^b \ge 0, \quad \forall i, \forall b, \forall t, \tag{5.11}$$

where $Q(y,u,v,g,\xi)$ is

$$\min \sum_t L_t s_t(\xi) \tag{5.12}$$

$$s.t. \quad s_t(\xi) \ge \sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b, \quad \forall t, \tag{5.13}$$

$$s_t(\xi) \ge \sum_b d_t^b - \sum_b \sum_i x_{it}^b - \sum_b w_t^b(\xi), \quad \forall t, \tag{5.14}$$

where, the first-stage objective function (5.1) consists of startup, shutdown and fuel costs. Constraints (5.2) and (5.3) force the minimum up-time and minimum down-time limits, respectively. Constraints (5.4) and (5.5) represent start-up and shut-down status constraints, respectively. Constraints (5.6) limit the generation capacity

lower and upper bounds. Constraints (5.7) and (5.8) represent the ramping constraints. Constraints (5.9) restrict the transmission line capacity limits. Chance constraints (5.10) ensure the chance that load imbalance violates a specified tolerance level is no more than a predefined risk level. Constraints (5.13) and (5.14) calculate the energy shortage or oversupply for each time period, respectively. In addition, we approximate the quadratic generation cost $F_i(.)$ using a J-piece piecewise linear function as below:

$$\phi_{it}^b \geq \alpha_{it}^{jb} y_{it}^b + \rho_{it}^{jb} x_{it}^b, \quad \forall t, \forall b, \forall i, j = 1, ..., J. \tag{5.15}$$

5.3.2 Confidence Set Construction

As discussed above, we allow the wind output distribution P to be unknown and ambiguous. We assume that P belongs to a confidence set $\mathbb{D}$ with a specific confidence level β (for example 95%). By utilizing the data information (i.e., historical data of wind output), we construct a distribution-based confidence set $\mathbb{D} = \{P \in \mathcal{M}_+ : d(P, \hat{P}) \leq \varphi\}$, where $\mathcal{M}_+$ is the set of all distributions, $\hat{P}$ represents the reference distribution, $d(P, \hat{P})$ is the probability distance between P and $\hat{P}$, and φ represents the tolerance level. We obtain the reference distribution by drawing a histogram using historical data. By partitioning the sample space Ω into N bins, such that $\Omega = \bigcup_{n=1}^{N} B_n$, and then counting the frequency of samples in each bin, i.e., S_n, we can obtain the reference distribution $\hat{P} = (\hat{p}_1, \hat{p}_2, ..., \hat{p}_N)$, in which $\hat{p}_n = S_n/S, \forall n$, and S is the total amount of data. To measure $d(P, \hat{P})$, in this study, we use L_∞ norm. Accordingly, we can define confidence set $\mathbb{D}$ as follows:

$$\mathbb{D} = \{P \in \mathbb{R}_+^N | \max_{1 \leq n \leq N} |p^n - \hat{p}^n| \leq \varphi\}. \tag{5.16}$$

According to [123], given S historical data points and N bins, the convergence rate between ambiguous distribution P and reference distribution $\hat{P}$ is as below:

$$Pr\{\|P - \hat{P}\|_\infty \leq \varphi\} \geq 1 - 2N \exp(-2S\varphi). \tag{5.17}$$

If the confidence level (the right-hand side of (5.17)) is equal to β, then, we have

$$\varphi = (1/2S)\log(2N/(1-\beta)). \tag{5.18}$$

From (5.18) we observe, as the number of historical observations increases to infinity, the value of φ goes to zero and $\hat{P}$ converges to P accordingly.

5.3.3 Data-Driven Chance Constraint and Its Reformulation

Since the true distribution of wind output is ambiguous within $\mathbb{D}$, we restrict that chance constraint (5.10) should be satisfied under the worst-case distribution in $\mathbb{D}$. Therefore, we can formulate the data-driven chance constraint as follows:

$$\min_{P\in\mathbb{D}} \quad Pr(-\pi \leq \sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b \leq \pi) \geq 1-\epsilon, \quad \forall t. \tag{5.19}$$

Inequality (5.19) can be further reformulated as

$$\min_{P\in\mathbb{D}} \quad \sum_{n=1}^{N} p_t^n . \mathbf{1}_{[-\pi,\pi]}(\sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi^n) - \sum_b d_t^b) \geq 1-\epsilon, \quad \forall t, \tag{5.20}$$

where, $\mathbf{1}_{[-\pi,\pi]}(.)$ is an indicator function that equals to 1 if $-\pi \leq \sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b \leq \pi$ and 0 otherwise, and ξ^n is the central point of bin n. Then, we let binary variables $z^n = \mathbf{1}_{[-\pi,\pi]}(\sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi^n) - \sum_b d_t^b)$; and use big-M method to reformulate (5.20) as the following MILP model:

$$\begin{aligned} -\pi - (1-z_t^n)M &\leq \sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b \\ &\leq (1-z_t^n)M + \pi, \quad \forall t, \forall n, \end{aligned} \tag{5.21}$$

$$\min_p \sum_{n=1}^{N} p_t^n z_t^n \geq 1-\epsilon, \quad \forall t, \tag{5.22}$$

$$\varphi + \hat{p}_t^n \geq p_t^n \geq -\varphi + \hat{p}_t^n, \quad \forall t, \forall n, \tag{5.23}$$

$$\sum_{n=1}^{N} p_t^n = 1, \quad \forall t, \tag{5.24}$$

$$p_t^n \geq 0, \quad \forall t, \forall n, \tag{5.25}$$

where constraints (5.23) and (5.24) are reformulation of the confidence set $\mathbb{D}$. To eliminate the minimization operation in constraint (5.22), we dualize (5.22) to (5.25) to get the following formulation:

$$\max \quad \sum_{n=1}^{N} \lambda_t^n(-\varphi + \hat{p}_t^n) - \tau_t^n(\varphi + \hat{p}_t^n) + \eta_t \tag{5.26}$$

$$s.t. \quad \eta_t + \lambda_t^n - \tau_t^n \leq z_t^n, \quad \forall t, \forall n, \tag{5.27}$$

$$\lambda_t^n, \tau_t^n \geq 0, \eta_t \text{ free}, \quad \forall t, \forall n, \tag{5.28}$$

where τ_t^n, λ_t^n and η_t are dual variables of constraints (5.23) and (5.24). Thus, constraint (5.22) to (5.25) can be reformulated as:

$$\sum_{n=1}^{N} \lambda_t^n(-\varphi + \hat{p}_t^n) - \tau_t^n(\varphi + \hat{p}_t^n) + \eta_t \geq 1 - \epsilon, \tag{5.29}$$

$$\text{Constraints (5.27)} - \text{(5.28)}. \tag{5.30}$$

5.3.4 Objective Reformulation

Based on the above-defined confidence set for the unknown wind output distribution, we formulate the data-driven chance-constrained two-stage UC model. In the second stage, we consider the worst-case imbalance penalty cost associated with the worst-case wind output distribution in $\mathbb{D}$. Since different scenarios ξ^n are independent, we can interchange the summation and the minimization operations. Hence, we can reformulate the second-stage objective function to its data-driven formulation.

$$\max_{P \in \mathbb{D}} E_P[Q(y, u, v, g, \xi)] = \max_{P \in \mathbb{D}} \min_{s} \sum_{n=1}^{N} \sum_{t} \; p_t^n \; L_t s_t(\xi^n). \tag{5.31}$$

5.4 Solution Methodology

We apply the Column-and-Constraint generation method [120], in a decomposition framework. We have the following master problem (MP):

$$\min_{y,u,v,\phi,g,z,\lambda,\tau,\eta} (SU_i^b u_{it}^b + SD_i^b v_{it}^b + \phi_{it}^b) + \vartheta \qquad \text{(MP)}$$

$$s.t. \quad \text{Constraints} \quad (5.2) - (5.8) \text{ and } (5.15)$$

$$\text{Constraints} \quad (5.9), (5.13), (5.14) \text{ and } (5.21), \ \forall n,$$

$$\text{Constraints} \quad (5.29), (5.27) \text{ and } (5.28),$$

$$\text{Optimality cuts,}$$

where ϑ is the second-stage optimal objective value. As for the subproblem (S), we initially dualize the second-stage problem, i.e., constraints (5.13) and (5.14). Then, we formulate S as the followings:

$$\omega(g) = \max_{\mu,\gamma,p} \sum_{n=1}^{N} \Big(\big(\sum_b \sum_i x_{it}^b + \sum_b w_t^b(\xi) - \sum_b d_t^b \big) \mu_t^n \qquad \text{(S)}$$

$$+ \big(\sum_b d_t^b - \sum_b \sum_i x_{it}^b - \sum_b w_t^b(\xi) \big) \gamma_t^n \Big) \tag{5.32}$$

$$s.t. \quad \mu_t^n \le L_t p_t^n, \quad \forall t, \forall n, \tag{5.33}$$

$$\gamma_t^n \le L_t p_t^n, \quad \forall t, \forall n, \tag{5.34}$$

$$-\varphi + \hat{p}_t^n \le p_t^n \le \varphi + \hat{p}_t^n, \quad \forall t, \forall n, \tag{5.35}$$

$$\sum_{n=1}^{N} p_t^n = 1, \quad \forall t, \tag{5.36}$$

$$\mu_t^n, \gamma_t^n, \hat{p}_t^n \ge 0, \quad \forall t, \forall n, \tag{5.37}$$

where constraints (5.35) to (5.36) represent the confidence set $\mathbb{D}$. $\hat{p}_t^n$ and p_t^n denote the reference distribution and the true distribution of wind output at time t for scenario n, respectively. Also, μ_t^n and γ_t^n are dual variables of constraints (5.13) and (5.14), respectively. We briefly describe the solution algorithm in the following steps:

1. Initialization. Set $k = 1$, $\vartheta = -\infty$

2. Solve MP and get the first-stage decision variables.

3. Fix the generation level g and solve S to obtain $\omega(g)$.

4. If $\omega(g) \leq \vartheta$, then stop and output the first-stage decisions. Otherwise, set $k = k + 1$. Generate and add the following cut (5.38) to MP and go to step 2.

$$\vartheta \geq \sum_{n=1}^{N} \sum_{t} p_t^n L_t s_t(\xi^n). \tag{5.38}$$

Notice here, since our model allows energy imbalance, the first-stage solutions are always feasible and no feasibility check is required in step 3.

5.5 Case Study

In order to test the effectivness of our proposed approach, we conduct experiments on a modified IEEE 118-bus system (available at http://motor.ece.iit.edu/data). To generate historical data set of wind output, we assume the uncertain wind power follows a multivariate normal distribution with mean equal to the forecasted value and variance equal to 0.3 of the mean. We set the number of bins to be 5. Then, using Monte Carlo simulation, we generate a historical data set for each wind farm in each time period. In addition, in the piecewise linear cost function, we set the number of pieces to be 5. We use C++ and CPLEX 12.6 to implement our model on a computer with Intel(R) Xeon(R) 3.2 and 8 GB memory.

5.5.1 Effects of the Historical Data

First, we show how the number of historical data points can effect the conservatism of the proposed approach. We set the confidence level β to be 99% and test the performance of the proposed DDCHC for different size of data ranging from 10 to 10000. We report the performance of the proposed approach in Table 5.1. We also report the result of the traditional chance-constrained UC (CCUC) for a set of 50000

historical data, as the benchmark. From Table 5.1 and Fig. 5.1, we can observe that as the size of data gets larger, both the value of φ and DDCHC objective value decrease. That is becuase, with larger data sets, $\mathbb{D}$ gets smaller and DDCHC becomes less conservative. Theoretically, as the size of historical data goes to infinity, the confidence set $\mathbb{D}$ shrinks (reference distribution converges to true distribution) and eventually DDCHC becomes risk neutral.

Table 5.1: Effects of historical data on total cost

# of	DDCHC			CCUC	
data	φ	Obj($m)	T(s)	Obj($m)	T(s)
10	0.34539	2.0886	12	1.5611	10
50	0.06908	1.6655	12	1.5611	10
100	0.03454	1.6130	11	1.5611	10
500	0.00691	1.5714	11	1.5611	10
1000	0.00345	1.5669	10	1.5611	10
5000	0.00069	1.5621	10	1.5611	10
10000	0.00035	1.5616	10	1.5611	10

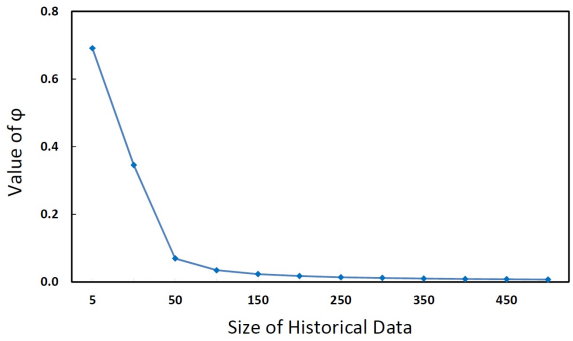

Figure 5.1: Effects of the size of historical data on the value of φ

5.5.2 Effects of the Confidence Level

In this section, we show by experiments that how the confidence level β effects the conservatism of DDCHC. We set the number of historical data to be 100 and test DDCHC performance over a range of confidence level between 0.5 to 0.99. Table 5.2 and Fig. 5.2 illustrate that as the value of β increases, both the value of φ and DDCHC objective value increase. In fact, larger φ means higher chance that $\mathbb{D}$ should include the true wind output distribution, and therefore $\mathbb{D}$ is larger. Hence, DDCHC becomes more conservative as the value of β increases.

Table 5.2: Effects of the confidence level on total cost

β	DDCHC	
	φ	Obj($m)
0.5	0.01498	1.5829
0.6	0.01609	1.5845
0.7	0.01753	1.5865
0.8	0.01956	1.5895
0.9	0.02303	1.5929
0.95	0.02649	1.6013
0.99	0.03454	1.6130

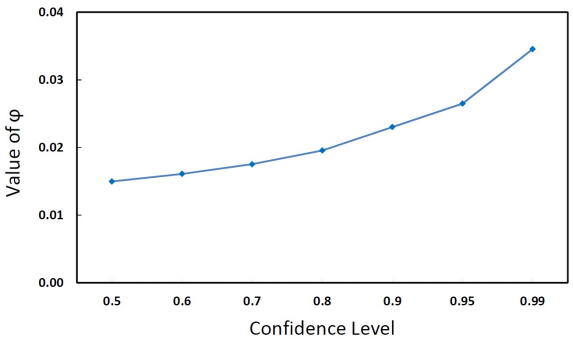

Figure 5.2: Effects of the confidence level on the value of φ

5.5.3 Comparison with Traditional Chance-Constrained UC

We compare the performance of the proposed DDCHC with that of CCUC in terms of the system reliability. For different size of historical data sets from 10 to 100, we first solve the modified 118-bus system by both DDCHC and CCUC. Then, we fix the optimal thermal generation levels obtained by DDCHC and CCUC and solve the second stage for random instances of wind output under randomly simulated distribution scenarios. We report the operational cost (OC), total cost (TC) and the computational time in Table 5.3. We observe that DDCHC results in higher operational costs but less total cost in comparison with CCUC. This is because, DDCHC is conservative, compared to CCUC, and schedules more generators such that they can accommodate the worst-case wind output distribution (to avoid large energy imbalances) for the whole next day. Therefore, DDCHC is a risk-averse approach. In fact, especially in case of unexpected realization of wind output, compared to CCUC operational decisions, DDCHC generation schedules and generation levels lead to less energy and load imbalances.

Table 5.3: DDCHC versus CCUC

# of	DDCHC			CCUC		
Data	OC ($m)	TC ($m)	T(s)	OC ($m)	TC ($m)	T(s)
10	1.1949	2.6937	12	1.1911	2.7678	9
20	1.1946	2.6872	12	1.1913	2.7603	9
30	1.1943	2.6807	12	1.1917	2.7555	9
50	1.1935	2.6723	12	1.1920	2.7494	9
100	1.1925	2.6644	11	1.1922	2.7413	9

5.6 Summary

In this chapter, we presented a data-driven chance-constrained stochastic unit commitment, in which the chance constraint controls the level of energy imbalance. Unlike traditional chance constraint which assumes that the unknown parameter follows a certain probability distribution, our proposed approach learns from a given historical data set to construct a confidence set for the unknown wind output distribution. Our model is a two-stage model which deals with the UC and ED decisions in the first stage and considers the worst-case energy imbalance cost associated with the worst-case wind output distribution in the second stage. By numerical experiments, we show our approach leads to more reliable and robust generation schedules than those of the traditional chance-constrained UC. We also show that as the number of historical data goes to infinity, the conservatism of the proposed approach vanishes eventually.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

Distributionally robust optimization (DRO) is a newly emerged approach to address optimization problems under uncertainty. By utilizing historical data and partial information of random parameter's probability distribution, DRO takes advantage of both stochastic programming and robust optimization approaches and bridges the gap between them and results in reliable but cost efficient decisions.

In this research, we addressed four major problems arising in power systems operations management. First, in chapter 2, we developed a DRO model for the transmission expansion planning problem under electricity demand and renewable energy uncertainties. Then, in chapter 3, we proposed a defender-attacker-defender based DRO approach for the transmission hardening planning of power systems with a high penetration of wind generation and random disruptive events. Afterward, in chapter 4, we developed two DRO models to evaluate the reliability of the power transmission system under $\mathrm{N}-k$ security criterion. Finally, in chapter 5, we formulated a DRO chance-constrained stochastic model for the power generation scheduling under renewable energy uncertainty. For all four problems, we successfully reformulated the original problems to two-stage stochastic mixed integer programs and solved them by decomposition algorithms. Our experimental results demonstrated the effectiveness of our DRO models compared to stochastic and robust optimization approaches.

In power generation scheduling under wind uncertainty, one recently suggested remedy to maintain the system reliability against high fluctuations of wind output is to utilize flexible resources, such as battery storages and quick-start generators.

However, by considering this type of resources, adjustable binary variables will appear in the second stage of the traditional two-stage UC formulation. For this case, the two-stage UC model can not be solved by the traditional decomposition algorithms. As a future work, we aim to develop a DRO two-stage UC with flexible resources and wind uncertainty and to solve it by a recently developed algorithm.

BIBLIOGRAPHY

[1] MM Adibi and LH Fink. Power system restoration planning. *IEEE Trans. Power Syst.*, 9(1):22–28, Feb. 1994.

[2] Natalia Alguacil, Andrés Delgadillo, and José M Arroyo. A trilevel programming approach for electric grid defense planning. *Comput. Oper. Res.*, 41:282–290, 2014.

[3] Natalia Alguacil, Alexis L Motto, and Antonio J Conejo. Transmission expansion planning: a mixed-integer LP approach. *IEEE Trans. Power Syst.*, 18(3):1070–1077, 2003.

[4] Behnam Alizadeh, Shahab Dehghan, Nima Amjady, Shahram Jadid, and Ahad Kazemi. Robust transmission system expansion considering planning uncertainties. *IET Gener. Transm. Distrib.*, 7(11):1318–1331, 2013.

[5] Fernando L Alvarado. Computational complexity in power systems. *IEEE Trans. Power App. Syst.*, 95(4):1028–1037, 1976.

[6] Göran Andersson, Peter Donalek, Richard Farmer, Nikos Hatziargyriou, Innocent Kamwa, Prabhashankar Kundur, Nelson Martins, John Paserba, Pouyan Pourbeik, Juan Sanchez-Gasca, R Schulz, Alex Stankovic, Carson Taylor, and Vijay Vittal. Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance. *IEEE Trans. Power Syst.*, 20(4):1922–1928, 2005.

[7] José Manuel Arroyo. Bilevel programming applied to power system vulnerability analysis under multiple contingencies. *IET Gener. Transm. Distrib.*, 4(2):178–190, 2010.

[8] José Manuel Arroyo, Natalia Alguacil, and Miguel Carrión. A risk-based approach for transmission network expansion planning under deliberate outages. *IEEE Trans. Power Syst.*, 25(3):1759–1766, 2010.

[9] Ali Bagheri, Jianhui Wang, and Chaoyue Zhao. Data-driven stochastic transmission expansion planning. *IEEE Trans. Power Syst.*, 32(5):3461–3470, 2017.

[10] Nagaraj Balijepalli, Subrahmanyam S Venkata, Charles W Richter, Richard D Christie, and Vito J Longo. Distribution system reliability assessment due to lightning storms. *IEEE Trans. Power Deliv.*, 20(3):2153–2159, 2005.

[11] Luis Baringo and Antonio J Conejo. Offering strategy via robust optimization. *IEEE Trans. Power Syst.*, 26(3):1418–1425, 2011.

[12] Rüdiger Barth, Heike Brand, Peter Meibom, and Christoph Weber. A stochastic unit-commitment model for the evaluation of the impacts of integration of large amounts of intermittent wind power. In *Proc. Int. Conf. Probabilistic Methods Applied to Power Syst.*, pages 1–8, 2006.

[13] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[14] Dimitris Bertsimas, Eugene Litvinov, Xu Andy Sun, Jinye Zhao, and Tongxin Zheng. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Trans. Power Syst.*, 28(1):52–63, 2013.

[15] Qiaoyan Bian, Huanhai Xin, Zhen Wang, Deqiang Gan, and Kit Po Wong. Distributionally robust solution to the reserve scheduling problem with partial information of wind power. *IEEE Trans. Power Syst.*, 30(5):2822–2823, 2015.

[16] Vicki M Bier, Eli R Gratz, Naraphorn J Haphuriwat, Wairimu Magua, and Kevin R Wierzbicki. Methodology for identifying near-optimal interdiction strategies for a power transmission system. *Reliab. Eng. Syst. Saf.*, 92(9):1155–1161, 2007.

[17] Roy Billinton and Ronald N Allan. *Reliability evaluation of power systems*. Springer Science & Business Media, 2013.

[18] Roy Billinton, Hua Chen, and Raymond Ghajar. A sequential simulation technique for adequacy evaluation of generating systems including wind energy. *IEEE Trans. Energy Conver.*, 11(4):728–734, 1996.

[19] Roy Billinton and Yi Gao. Multistate wind energy conversion system models for adequacy assessment of generating systems incorporating wind energy. *IEEE Trans. Energy Conver.*, 23(1):163–170, 2008.

[20] Roy Billinton and A Jonnavithula. Composite system adequacy assessment using sequential Monte Carlo simulation with variance reduction techniques. *Proc. Inst. Elect. Eng., Gener., Transm., Distrib.*, 144(1):1–6, 1997.

[21] Severin Borenstein and James Bushnell. The US electricity industry after 20 years of restructuring. *Annu. Rev. Econ.*, 7(1):437–463, 2015.

[22] FranÇois Bouffard and Francisco D Galiana. Stochastic security for operations planning with significant wind power generation. *IEEE Trans. Power Syst.*, 23(2):306–316, 2008.

[23] Gerald Brown, Matthew Carlyle, Javier Salmerón, and Kevin Wood. Analyzing the vulnerability of critical infrastructure to attack and planning defenses. *Tutorials in Oper. Res.: Emerging Theory, Methods, and Applications*, pages 102–123, 2005.

[24] Gerald Brown, Matthew Carlyle, Javier Salmerón, and Kevin Wood. Defending critical infrastructure. *Interfaces*, 36(6):530–544, 2006.

[25] M Oloomi Buygi, M Shahidehpour, H Modir Shanechi, and G Balzer. Market based transmission planning under uncertainties. In *Proc. Int. Conf. Probabilistic Methods Applied to Power Syst.*, pages 563–568, 2004.

[26] Majid Oloomi Buygi, Gerd Balzer, Hasan Modir Shanechi, and Mohammad Shahidehpour. Market-based transmission expansion planning. *IEEE Trans. Power Syst.*, 19(4):2060–2067, 2004.

[27] Majid Oloomi Buygi, Hasan Modir Shanechi, Gerd Balzer, Mohammad Shahidehpour, and Nasser Pariz. Network planning in unbundled power systems. *IEEE Trans. Power Syst.*, 21(3):1379–1387, 2006.

[28] Bruno Canizes, João Soares, Zita Vale, and HM Khodr. Hybrid fuzzy Monte Carlo technique for reliability assessment in transmission power systems. *Energy*, 45(1):1007–1017, 2012.

[29] Bokan Chen, Jianhui Wang, Lizhi Wang, Yanyi He, and Zhaoyu Wang. Robust optimization for transmission expansion planning: Minimax cost vs. minimax regret. *IEEE Trans. Power Syst.*, 29(6):3069–3077, 2014.

[30] Qiming Chen, Chuanwen Jiang, Wenzheng Qiu, and James D McCalley. Probability models for estimating the probabilities of cascading outages in high-voltage transmission network. *IEEE Trans. Power Syst.*, 21(3):1423–1431, 2006.

[31] Qiming Chen and James D McCalley. Identifying high risk N-k contingencies for online security assessment. *IEEE Trans. Power Syst.*, 20(2):823–834, 2005.

[32] Jaeseok Choi, A El-Keib, and Trungtinh Tran. A fuzzy branch and bound-based transmission system expansion planning for the highest satisfaction level of the decision maker. *IEEE Trans. Power Syst.*, 20(1):476–484, 2005.

[33] Antonio J Conejo, Miguel Carrión, and Juan M Morales. *Decision making under uncertainty in electricity markets*, volume 1. Springer, 2010.

[34] Armando M Leite da Silva, Leonidas Chaves de Resende, Luiz Antônio da Fonseca Manso, and Vladimiro Miranda. Composite reliability assessment based on Monte Carlo simulation and artificial neural networks. *IEEE Trans. Power Syst.*, 22(3):1202–1209, 2007.

[35] C Matthew Davis and Thomas J Overbye. Multiple element contingency screening. *IEEE Trans. Power Syst.*, 26(3):1294–1301, 2011.

[36] Sebastián de la Torre, Antonio J Conejo, and Javier Contreras. Transmission expansion planning in electricity markets. *IEEE Trans. Power Syst.*, 23(1):238–248, 2008.

[37] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.*, 58(3):595–612, 2010.

[38] Ronan Doherty and Mark O'malley. A new approach to quantify reserve demand in systems with significant installed wind capacity. *IEEE Trans. Power Syst.*, 20(2):587–595, 2005.

[39] Yury Dvorkin, Hrvoje Pandžić, Miguel A Ortega-Vazquez, and Daniel S Kirschen. A hybrid stochastic/interval approach to transmission-constrained unit commitment. *IEEE Trans. Power Syst.*, 30(2):621–631, 2015.

[40] Margaret J Eppstein and Paul DH Hines. A random chemistry algorithm for identifying collections of multiple contingencies that initiate cascading failure. *IEEE Trans. Power Syst.*, 27(3):1698–1705, 2012.

[41] Lei Fan, Jianhui Wang, Ruiwei Jiang, and Yongpei Guan. Min-max regret bidding strategy for thermal generator considering price uncertainty. *IEEE Trans. Power Syst.*, 29(5):2169–2179, 2014.

[42] Neng Fan, David Izraelevitz, Feng Pan, Panos M Pardalos, and Jianhui Wang. A mixed integer programming approach for optimal power grid intentional islanding. *Energ. Syst.*, 3(1):77–93, 2012.

[43] Risheng Fang and David J Hill. A new strategy for transmission expansion in competitive electricity markets. *IEEE Trans. Power Syst.*, 18(1):374–380, 2003.

[44] Bruno Fanzeres, Alexandre Street, and Luiz Augusto Barroso. Contracting strategies for renewable generators: a hybrid stochastic and robust optimization approach. *IEEE Trans. Power Syst.*, 30(4):1825–1837, 2015.

[45] Alison L Gibbs and Francis Edward Su. On choosing and bounding probability metrics. *Int. Stat. Rev.*, 70(3):419–435, 2002.

[46] Joel Goh and Melvyn Sim. Distributionally robust optimization and its tractable approximations. *Oper. Res.*, 58(4):902–917, 2010.

[47] Andrés D González, Leonardo Dueñas-Osorio, Mauricio Sánchez-Silva, and Andrés L Medaglia. The interdependent network design problem for optimal infrastructure system restoration. *Comput.-Aided Civ. Inf.*, 31(5):334–350, 2016.

[48] Reinaldo A González-Fernández, Armando M Leite da Silva, Leonidas C Resende, and Marcus T Schilling. Composite systems reliability evaluation based on Monte Carlo simulation and cross-entropy methods. *IEEE Trans. Power Syst.*, 28(4):4598–4606, 2013.

[49] Cliff Grigg, Peter Wong, Paul Albrecht, Ron Allan, Murty Bhavaraju, Roy Billinton, Quan Chen, Clement Fong, Suheil Haddad, Sastry Kuruganty, W Li, R Mukerji, D Patton, N Rau, D Reppen, A Schneider, Mohammad Shahidehpour, and C Singh. The IEEE reliability test system-1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Trans. Power Syst.*, 14(3):1010–1020, 1999.

[50] Kory W Hedman, Michael C Ferris, Richard P O'Neill, Emily Bartholomew Fisher, and Shmuel S Oren. Co-optimization of generation unit commitment and transmission switching with N-1 reliability. *IEEE Trans. Power Syst.*, 25(2):1052–1063, 2010.

[51] Eitan Israeli. *System interdiction and defense*. PhD thesis, Monterey, California. Naval Postgraduate School, 1999.

[52] Rabih A Jabr. Robust transmission network expansion planning with uncertain renewable generation and loads. *IEEE Trans. Power Syst.*, 28(4):4558–4567, 2013.

[53] Youwei Jia, Ke Meng, and Zhao Xu. N-k induced cascading contingency screening. *IEEE Trans. Power Syst.*, 30(5):2824–2825, 2015.

[54] Youwei Jia, Zhao Xu, Loi Lei Lai, and Kit Po Wong. Risk-based power system security analysis considering cascading outages. *IEEE Trans. Ind. Informat.*, 12(2):872–882, 2016.

[55] Ruiwei Jiang and Yongpei Guan. Data-driven chance constrained stochastic program. *Math. Program.*, 158(1):291–327, 2016.

[56] Ruiwei Jiang, Jianhui Wang, and Yongpei Guan. Robust unit commitment with wind power and pumped storage hydro. *IEEE Trans. Power Syst.*, 27(2):800–810, 2012.

[57] Ruiwei Jiang, Muhong Zhang, Guang Li, and Yongpei Guan. Two-stage robust power grid optimization problem. *submitted to Journal of Operations Research*, 2010.

[58] Panida Jirutitijaroen and Chanan Singh. Reliability constrained multi-area adequacy planning using stochastic programming with sample-average approximations. *IEEE Trans. Power Syst.*, 23(2):504–513, 2008.

[59] Aleksandra Kanevce, Igor Mishkovski, and Ljupco Kocarev. Modeling long-term dynamical evolution of Southeast European power transmission system. *Energy*, 57:116–124, 2013.

[60] P Kaplunovich and K Turitsyn. Fast and reliable screening of N-2 contingencies. *IEEE Trans. Power Syst.*, 31(6):4243–4252, 2016.

[61] Hagkwen Kim and Chanan Singh. Reliability modeling and simulation in power systems with aging characteristics. *IEEE Trans. Power Syst.*, 25(1):21–28, 2010.

[62] Prabha Kundur, John Paserba, Venkat Ajjarapu, Göran Andersson, Anjan Bose, Claudio Canizares, Nikos Hatziargyriou, David Hill, Alex Stankovic, Carson Taylor, Thierry Van Cutsem, and Vijay Vittal. Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Trans. Power Syst.*, 19(3):1387–1401, 2004.

[63] Kristina Hamachi LaCommare and Joseph H Eto. Cost of power interruptions to electricity consumers in the United States (US). *Energy*, 31(12):1845–1855, 2006.

[64] W Li, X Xiong, and J Zhou. Incorporating fuzzy weather-related outages in transmission system reliability assessment. *IET Gener. Transm. Distrib.*, 3(1):26–37, 2009.

[65] Wenyuan Li. Incorporating aging failures in power system reliability evaluation. *IEEE Trans. Power Syst.*, 17(3):918–923, 2002.

[66] Wenyuan Li and Jiping Lu. Risk evaluation of combinative transmission network and substation configurations and its application in substation planning. *IEEE Trans. Power Syst.*, 20(2):1144–1150, 2005.

[67] D Lindenmeyer, HW Dommel, and MM Adibi. Power system restoration-a bibliographical survey. *Int. J. Elec. Power*, 23(3):219–227, 2001.

[68] Juan Álvarez López, Kumaraswamy Ponnambalam, and Víctor H Quintana. Generation and transmission expansion under risk using stochastic programming. *IEEE Trans. Power Syst.*, 22(3):1369–1378, 2007.

[69] Executive Office of the President. Economic benefits of increasing electric grid resilience to weather outages. 2013.

[70] NERC reliability standard. PRC-023-3. Aug. 2014.

[71] North Electric Reliability Corporation. NERC operating manual. 2016.

[72] PAC (Protection, Automation, Control) Magazine.

[73] U.S. Department of Energy. 20% wind energy by 2030: Increasing wind energy's contribution to U.S. electricity supply. 2008.

[74] U.S. Department of Energy. Hardening and resiliency: US energy industry response to recent hurricane seasons. 2010.

[75] US Energy Information Administration. Annual energy outlook. 2013.

[76] U.S. Energy Information Administration. Electric power monthly. 2017.

[77] Lamine Mili, Q Qiu, and Arun G Phadke. Risk assessment of catastrophic failures in electric power systems. *Int. J. Crit. Infrastruct.*, 1(1):38–63, 2004.

[78] Alberto Moreira, Alexandre Street, and Jose Manuel Arroyo. An adjustable robust optimization approach for contingency-constrained transmission expansion planning. *IEEE Trans. Power Syst.*, 30(4):2013–2022, 2015.

[79] Alexis L Motto, José M Arroyo, and Francisco D Galiana. A mixed-integer LP procedure for the analysis of electric grid security under disruptive threat. *IEEE Trans. Power Syst.*, 20(3):1357–1365, 2005.

[80] Miguel A Ortega-Vazquez and Daniel S Kirschen. Estimating the spinning reserve requirements in systems with significant wind power generation penetration. *IEEE Trans. Power Syst.*, 24(1):114–124, 2009.

[81] U Aytun Ozturk, Mainak Mazumdar, and Bryan A Norman. A solution to the stochastic unit commitment problem using chance constrained programming. *IEEE Trans. Power Syst.*, 19(3):1589–1598, 2004.

[82] Hrvoje Pandžić, Yury Dvorkin, Ting Qiu, Yishen Wang, and Daniel S Kirschen. Toward cost-efficient and reliable unit commitment under uncertainty. *IEEE Trans. Power Syst.*, 31(2):970–982, 2016.

[83] Anthony Papavasiliou, Shmuel S Oren, and Richard P O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Trans. Power Syst.*, 26(4):2197–2206, 2011.

[84] Mario VF Pereira and Neal J Balu. Composite generation/transmission reliability evaluation. *Proceedings of the IEEE*, 80(4):470–491, 1992.

[85] Georg Pflug and David Wozabal. Ambiguity in portfolio selection. *Quant. Financ.*, 7(4):435–442, 2007.

[86] Naran M Pindoriya, Panida Jirutitijaroen, Dipti Srinivasan, and Chanan Singh. Composite reliability evaluation using Monte Carlo simulation and least squares support vector classifier. *IEEE Trans. Power Syst.*, 26(4):2483–2490, 2011.

[87] Svetlozar T Rachev and Ludger Rüschendorf. *Mass Transportation Problems: Volume I: Theory*, volume 1. Springer Science & Business Media, 1998.

[88] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *J. Risk*, 2:21–42, 2000.

[89] Jae Hyung Roh, Mohammad Shahidehpour, and Lei Wu. Market-based generation and transmission planning with uncertainties. *IEEE Trans. Power Syst.*, 24(3):1587–1598, 2009.

[90] Natalia Romero, Linda K Nozick, Ian Dobson, Ningxiong Xu, and Dean A Jones. Seismic retrofit for electric power systems. *Earthq. Spectra*, 31(2):1157–1176, 2015.

[91] Pablo A Ruiz, C Russ Philbrick, Eugene Zak, Kwok W Cheung, and Peter W Sauer. Uncertainty management in the unit commitment problem. *IEEE Trans. Power Syst.*, 24(2):642–651, 2009.

[92] S Ryan, James McCalley, and D Woodruff. Long term resource planning for electric power systems under uncertainty. Technical report, Iowa State Univ, 2011.

[93] Javier Salmeron, Kevin Wood, and Ross Baldick. Analysis of electric grid security under terrorist threat. *IEEE Trans. Power Syst.*, 19(2):905–912, 2004.

[94] Javier Salmeron, Kevin Wood, and Ross Baldick. Worst-case interdiction analysis of large-scale electric power grids. *IEEE Trans. Power Syst.*, 24(1):96–104, 2009.

[95] Mohammad Shahidehpour, F Tinney, and Yong Fu. Impact of security on power systems operation. *Proc. IEEE*, 93(11):2013–2025, 2005.

[96] Mohammad Shahidehpour, Hatim Yamin, and Zuyi Li. *Frontmatter and index*. Wiley Online Library, 2002.

[97] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2009.

[98] Zhen Shu and Panida Jirutitijaroen. Latin hypercube sampling techniques for power systems reliability analysis with renewable energy sources. *IEEE Trans. Power Syst.*, 26(4):2066–2073, 2011.

[99] Adam B Smith and Richard W Katz. US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Natural hazards*, 67(2):387–410, 2013.

[100] Wei Sun, Chen-Ching Liu, and Li Zhang. Optimal generator start-up strategy for bulk power system restoration. *IEEE Trans. Power Syst.*, 26(3):1357–1366, Aug. 2011.

[101] Alexandre B Tsybakov. *Introduction to nonparametric estimation*. Springer Science & Business Media, 2008.

[102] Aidan Tuohy, Eleanor Denny, and Mark O Malley. Rolling unit commitment for systems with significant installed wind capacity. In *Proc. 2007 IEEE Lausanne Power Tech*, pages 1380–1385, 2007.

[103] Aidan Tuohy, Peter Meibom, Eleanor Denny, and Mark O'Malley. Unit commitment for systems with significant wind penetration. *IEEE Trans. Power Syst.*, 24(2):592–601, 2009.

[104] Bart C Ummels, Madeleine Gibescu, Engbert Pelgrum, Wil L Kling, and Arno J Brand. Impacts of wind power on thermal generation unit commitment and dispatch. *IEEE Trans. energy conver.*, 22(1):44–51, 2007.

[105] Marianna Vaiman, Keith Bell, Yousu Chen, Badrul Chowdhury, Ian Dobson, Paul Hines, Milorad Papic, Stephen Miller, and Pei Zhang. Risk assessment of cascading outages: Methodologies and challenges. *IEEE Trans. Power Syst.*, 27(2):631–641, May 2012.

[106] Jianhui Wang, Audun Botterud, Vladimiro Miranda, Cláudio Monteiro, and Gerald Sheble. Impact of wind power forecasting on unit commitment and dispatch. In *Proc. 8th Int. Workshop Large-Scale Integration of Wind Power into Power Systems*, 2009.

[107] Jianhui Wang, Mohammad Shahidehpour, and Zuyi Li. Security-constrained unit commitment with volatile wind power generation. *IEEE Trans. Power Syst.*, 23(3):1319–1327, 2008.

[108] Qianfan Wang, Yongpei Guan, and Jianhui Wang. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Trans. Power Syst.*, 27(1):206–215, 2012.

[109] Yezhou Wang, Chen Chen, Jianhui Wang, and Ross Baldick. Research on resilience of power systems under natural disasters –A review. *IEEE Trans. Power Syst.*, 31(2):1604–1613, 2016.

[110] Zhen Wang, Qiaoyan Bian, Huanhai Xin, and Deqiang Gan. A distributionally robust co-ordinated reserve scheduling model considering CVaR-based wind power reserve requirements. *IEEE Trans. Sustain. Energy*, 7(2):625–636, 2016.

[111] Wei Wei, Feng Liu, and Shengwei Mei. Distributionally robust co-optimization of energy and reserve dispatch. *IEEE Trans. Power Syst.*, 7(1):289–300, 2016.

[112] Allen J Wood and Bruce F Wollenberg. *Power generation, operation, and control*. John Wiley & Sons, 2012.

[113] David Wozabal. A framework for optimization under ambiguity. *Ann. Oper. Res.*, 193(1):21–47, 2012.

[114] Yafei Yang, Xiaohong Guan, and Qiaozhu Zhai. Fast grid security assessment with N-k contingencies. *IEEE Trans. Power Syst.*, 2016.

[115] Yiming Yao, Thomas Edmunds, Dimitri Papageorgiou, and Rogelio Alvarez. Trilevel optimization in power network defense. *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, 37(4):712–718, 2007.

[116] H Yu, CY Chung, KP Wong, and JH Zhang. A chance constrained transmission network expansion planning method with consideration of load and wind farm uncertainties. *IEEE Trans. Power Syst.*, 24(3):1568–1576, 2009.

[117] Wei Yuan, Jianhui Wang, Feng Qiu, Chen Chen, Chongqing Kang, and Bo Zeng. Robust optimization-based resilient distribution network planning against natural disasters. *IEEE Trans. Smart Grid*, 7(6):2817–2826, 2016.

[118] Wei Yuan, Long Zhao, and Bo Zeng. Optimal power grid protection through a defender–attacker–defender model. *Reliab. Eng. Syst. Saf.*, 121:83–89, 2014.

[119] LIU Yutian, FAN Rui, and Vladimir Terzija. Power system restoration: a literature review from 2006 to 2016. *J. Mod. Power Syst. Cle.*, 4(3):332–341, 2016.

[120] Bo Zeng and Long Zhao. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Oper. Res. Lett.*, 41(5):457–461, 2013.

[121] Yiling Zhang, Siqian Shen, and Johanna L Mathieu. Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. *IEEE Trans. Power Syst.*, 32(2):1378–1388, 2017.

[122] Chaoyue Zhao and Yongpei Guan. Unified stochastic and robust unit commitment. *IEEE Trans. Power Syst.*, 28(3):3353–3361, 2013.

[123] Chaoyue Zhao and Yongpei Guan. Risk-averse data-driven optimization approach on solving the facility location problem. Technical report, 2014.

[124] Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse stochastic optimization with Wasserstein metric. 2015.

[125] Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse two-stage stochastic program with ζ-structure probability metrics. 2015.

[126] Chaoyue Zhao and Yongpei Guan. Data-driven stochastic unit commitment for integrating wind generation. *IEEE Trans. Power Syst.*, 31(4):2587–2596, 2016.

[127] Chaoyue Zhao and Ruiwei Jiang. Distributionally robust contingency-constrained unit commitment. *IEEE Trans. Power Syst.*, 2017.

[128] Chaoyue Zhao, Jianhui Wang, Jean-Paul Watson, and Yongpei Guan. Multi-stage robust unit commitment considering wind and demand response uncertainties. *IEEE Trans. Power Syst.*, 28(3):2708–2717, 2013.

[129] Chaoyue Zhao, Qianfan Wang, Jianhui Wang, and Yongpei Guan. Expected value and chance constrained stochastic unit commitment ensuring wind power utilization. *IEEE Trans. Power Syst.*, 29(6):2696–2705, 2014.

[130] Bie Zhaohong and Wang Xifan. Studies on variance reduction technique of Monte Carlo simulation in composite system reliability evaluation. *Electr. Power Syst. Res.*, 63(1):59–64, 2002.

[131] Steve Zymler, Daniel Kuhn, and Berç Rustem. Distributionally robust joint chance constraints with second-order moment information. *Math. Program.*, 137(1):167–198, 2013.

APPENDIX A

PROOF OF PROPOSITION 1

In order to prove Proposition 1, we prove the following two lemmas first.

Lemma A.1 *Given a set of historical data of size S and N bins, we have*

$$Pr(|\mathrm{p}^n - \hat{\mathrm{p}}^n| \geq \delta) \leq 2\exp(-2S\delta), \forall n = 1, ..., N. \tag{A.1}$$

Proof. Let y_s^n denote whether the observation s falls in the bin n, then $y_s^n \sim \text{Bernoulli}(\mathrm{p}^n)$, $\forall n = 1, ..., N$, $\forall s = 1, ..., S$. We also have $\hat{\mathrm{p}}^n = \sum_{s=1}^{S} y_s^n/S$ and $E[\hat{\mathrm{p}}^n] = \mathrm{p}^n$. Then, according to Hoeffding's Inequality for Bernoulli random variables, (A.1) holds. ■

In order to get the convergence rate under Wasserstein metric, we will study another metric, i.e., Total Variance metric, first, and by taking advantage of the relationship between two metrics, we will obtain the convergence rate of the Wasserstein metric. According to [45], for space $\mathcal{W}$, the Total Variation metric is defined as $d_{\mathrm{TV}}(\mathrm{P}, \hat{\mathrm{P}}) = \sum_{n=1}^{N} |\mathrm{p}^n - \hat{\mathrm{p}}^n|$.

Lemma A.2 *Given a set of historical data of size S and N bins, under Total Variation metric we have*

$$Pr(d_{\mathrm{TV}}(\mathrm{P}, \hat{\mathrm{P}}) \geq \delta) \leq 2N\exp(-2S\delta/N). \tag{A.2}$$

Proof. Due to the fact that $Pr(\sum_{n=1}^{N} |\mathrm{p}^n - \hat{\mathrm{p}}^n| < \delta) \geq Pr(\cap_{n=1}^{N}[|\mathrm{p}^n - \hat{\mathrm{p}}^n| < \delta/N])$,

we have

$$\begin{aligned} Pr(d_{\mathrm{TV}}(\mathrm{P},\hat{\mathrm{P}}) \geq \delta) &= Pr(\sum_{n=1}^{N} |\mathrm{p}^n - \hat{\mathrm{p}}^n| \geq \delta) \\ &\leq Pr(\cup_{n=1}^{N} [|\mathrm{p}^n - \hat{\mathrm{p}}^n| \geq \delta/N]) \\ &\leq \sum_{n=1}^{N} Pr([|\mathrm{p}^n - \hat{\mathrm{p}}^n| \geq \delta/N]) \\ &\leq 2N \exp(-2S\delta/N). \end{aligned} \tag{A.3}$$

where the inequality (A.3) holds due to Lemma 1. ■

Based on the above two lemmas, we are ready to prove Proposition 1. First, according to [45], we have the following relationship between the Wasserstein metric and the Total Variation metric:

$$d_{\mathrm{w}}(\mathrm{P},\hat{\mathrm{P}}) \leq \frac{D}{2}\, d_{\mathrm{TV}}(\mathrm{P},\hat{\mathrm{P}}), \tag{A.4}$$

where D denoted the diameter of the supporting space $\mathcal{W}$. Moreover, according to Lemma 2, we have

$$Pr(d_{\mathrm{TV}}(\mathrm{P},\hat{\mathrm{P}}) \leq \delta) \geq 1 - 2N \exp(-2S\delta/N). \tag{A.5}$$

Inequality (A.4) implies that $\frac{2d_{\mathrm{w}}(\mathrm{P},\hat{\mathrm{P}})}{D} \leq d_{\mathrm{TV}}(\mathrm{P},\hat{\mathrm{P}})$. Therefore,

$$\begin{aligned} Pr(\frac{2d_{\mathrm{w}}(\mathrm{P},\hat{\mathrm{P}})}{D} \leq \delta) \quad &\geq \quad Pr(d_{\mathrm{TV}}(\mathrm{P},\hat{\mathrm{P}}) \leq \delta) \\ &\geq 1 - 2N \exp(-2S\delta/N). \end{aligned} \tag{A.6}$$

Let $\varphi = D\delta/2$, then the proof of Proposition 1 is done.

VITA

Ali Bagheri

Candidate for the Degree of

Doctor of Philosophy

Dissertation: DATA-DRIVEN OPTIMIZATION IN POWER SYSTEMS OPERATIONS

Major Field: Industrial Engineering and Management

Biographical:

Personal Data: Born in Tehran, Iran

Education:

- Received the B.S. degree from Yazd University, Yazd, Iran, 2005, in Applied Mathematics.
- Received the M.S. degree from Mazandaran University of Science and Technology, Babol, Iran, 2008, in Industrial Engineering.
- Completed the requirements for the degree of Doctor of Philosophy with a major in Industrial Engineering and Management at Oklahoma State University, 2018.

Selected Publications:

- Bagheri, A., Zhao, C., and Guo, Y. Data-driven chance-constrained stochastic unit commitment under wind power uncertainty. IEEE Power & Energy Society General Meeting, Chicago, IL, 2017.
- A. Bagheri, J. Wang, C. Zhao, Data-driven stochastic transmission expansion planning, IEEE Transactions on Power Systems, 32 (5), 3461-3470, 2017.
- A. Bagheri, J. Wang, F. Qiu, C. Zhao. Resilient transmission hardening planning in a high renewable penetration era, accepted for publication in IEEE Transactions on Power Systems.
- A. Bagheri, C. Zhao. Distributionally robust transmission system reliability analysis under $\mathrm{N}-k$ security criterion, under the second review at IEEE Transactions on Reliability.